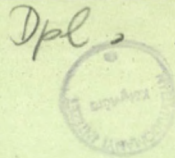


Évk.: 95. I. 16.



**Tanulmányok**  
**az elméleti és az alkalmazott fonetika köréből**

# **BESZÉDKUTATÁS**

**1993**







# **BESZÉDKUTATÁS**

Tanulmányok az elméleti  
és az alkalmazott fonetika köréből







# **BESZÉDKUTATÁS**

Tanulmányok az elméleti  
és az alkalmazott fonetika köréből

Szerkesztette: Gósy Mária és Siptár Péter

A Magyar Tudományos Akadémia Nyelvtudományi Intézete

Budapest 1993



ISBN 963 8461 62 1

© Az MTA Nyelvtudományi Intézete

Felelős kiadó: Kiefer Ferenc



## TARTALOM

Előszó (Gósy Mária – Siptár Péter)

Megnyitó (Vértés O. András)

**Siptár Péter:**

A Lexikális Fonológia elméletéről..... 1

**Gósy Mária:**

A lexikális hozzáférés (szófelismerési stratégiák) ..... 14

**Fagyal Zsuzsa:**

A magyar beszéd nyomainak percepciós vizsgálata a francia nyelvben..... 33

**Wacha Imre:**

A szövegfonetikai kutatások gondjairól..... 49

**Kassai Ilona:**

Gyorsult-e a magyar beszéd tempója? ..... 62

**Papp Ferenc:**

Személyi számítógép és fonetikai-fonológiai kutatások ..... 70

**Tarnóczy Tamás:**

Redukált fonetikai rendszer a gépi beszédfelismerés megkönnyítésére ..... 80

**Olaszy Gábor:**

Hangidőtartamok számítógépes elemzése a beszéd ritmikai szerkezetének vizsgálatához ..... 116

**Vicsi Klára – Vassné Kovács Emőke – Barczikai Péter:**

Beszédképzést segítő vizuális (gépi) módszer ..... 128

**Hegy Ágnes:**

Nyelvi struktúrák helyreállítása globális afáziás betegek kezelésében..... 144

**Balázs Boglárka:**

Hangképzési zavarok időskorban..... 156

**Molnár Sándor – Tatai Péter:**

Szubjektív és objektív beszédminősítési módszerek..... 166

**Molnár Ildikó:**

Tévesztés és megakadás kötött szöveg felolvasásában..... 175

**Laczkó Mária:**

A tempó és a szünet viszonya a hangos olvasásban..... 185

**A. Jászó Anna:**

A fonetikai kutatás újabb eredményeinek és a kiejtés tanításának kapcsolata..... 194

## ELŐSZÓ

Az MTA Nyelvtudományi Intézetének Fonetikai Laboratóriuma 1992 őszén BESZÉDKUTATÁS '92 főcímmel kétnapos országos konferenciát szervezett Budapesten. A konferencián az elméleti és alkalmazott fonetika területéről hangzottak el előadások nagyszámú közönség részvételével. A cél egyfelől a beszédkutatás szintjének, illetőleg a gyakorlati vonatkozásoknak a bemutatása volt a szakmai és az érdeklődő hallgatóság számára; másfelől annak demonstrálása, hogy a 3 évvel korábban ugyanitt megrendezett nemzetközi SPEECH RESEARCH '89 konferencia óta milyen új eredmények születtek ezen a területen. Tekintettel arra, hogy az akkori konferencia előadásai angol és német nyelven hangzottak el (vö. Proceedings of the Speech Research '89 International Conference, Hungarian Papers in Phonetics 21. 1989; Additional Papers, Hungarian Papers in Phonetics 22. 1990), egyre nagyobb igény mutatkozott egy magyar nyelvű találkozóra.

A BESZÉDKUTATÁS '92 szakmai sikere, az előadások színvonala, a széleskörű érdeklődés az egyes témák iránt vezetett e tanulmánykötet létrehozásának gondolatáig. Nem konferenciakötetet akartunk szerkeszteni, hanem egy olyan összeállítást, amely tematikájában ugyan megfelel a konferencián elhangzottaknak, de a szerzőknek lehetőségük volt az ott elhangzott megjegyzések, viták alapján az előadás anyagát minőségileg és mennyiségileg is átdolgozni.

E kötet válogatás, nem tartalmazza a kétnapos konferencia valamennyi előadását és témakörét. A szűkítésre a jelen könyvkiadás gazdasági helyzete éppúgy okot adott, mint az a természetes igény, hogy nyomtatásban csak a jó színvonalú, illetőleg újdonságot tartalmazó tanulmányok jelenjenek meg. Mindennek ellenére a kötet egyes tanulmányai a témára fordított kutatási idő és a kidolgozottság szintje tekintetében különbözőek. Ezt a fajta "heterogenitást" a szerkesztők annak reményében vállalták, hogy a kötet egésze, a magyar beszédkutatásról alkotott kép, ezt ellensúlyozni fogja.

A szerzők többsége fonetikus, nyelvész, a beszédkutatás más területeiről pedig fizikusok, mérnökök, logopédusok, foniáterek, azaz mindazok, akik ezen az interdiszciplináris területen dolgoznak. A témák a fonológiától az intonációkutatáson át a percepcióig a legkülönbözőbb kutatási problémákat, oktatási, társadalmi, orvosi és ipari alkalmazásokat érintik.

A BESZÉDKUTATÁS konferenciát VÉRTES O. ANDRÁS, a magyar fonetika legnagyobb élő reprezentánsa nyitotta meg. A következőkben az ő akkori gondolatai alapozzák meg a jelen tanulmánykötetet is.

*A szerkesztők*



## MEGNYITÓ

### Hölgyeim és uraim! Kedves Barátaim!

Konferenciánk alcíme: "Nyelvészeti és alkalmazott fonetika".

A hangtannal és a hangtan alkalmazásával kapcsolatban néhány gyakorlati, illetőleg elméleti kérdést érintenék.

**1. Fonetika és alkalmazott fonetika.** – A fonetika gyakorlati alkalmazásait tárgyaló munkára könnyen rá lehet mondani, hogy nem fonetikai, hiszen e tudománynak pusztán alkalmazásairól van szó. De a bíráló lehet éppen ellenkező irányú is: ha a tudományos fokozatra pályázó jelölt a fonetikának elméleti kérdéseit tárgyalja, akkor azt a véleményt nyilváníthatja a bíráló: hogy "szép, szép, de mi a disszertációnak a társadalmi hasznossága?"

Nézetem szerint egy tudományt és a részben vagy nagyrészt erre épülő alkalmazott tudományt nem szabad mereven elválasztani egymástól. A kettőt lehet egymástól elhatárolva tárgyalni, de ha a kutatás szempontjai úgy kívánják, a kettőt egy nagyobb egységben is fölfoghatjuk.

Kétnapos konferenciánkat e nagyobb egység szellemében tartjuk.

**2. A társadalmi tudományok és a természettudományok egymást megtermékenyítő hatásáról.** – Nem is olyan régen egy kitűnő nyelvtudományi kandidátusi értekezés vitájában a bíráló bizottság egyik tagja azt a nézetét (vagy sejtését?) fejezte ki, hogy hiábavaló igyekezet a társadalmi (németes szóval: "humán") tudományok művelőinek a természettudományokra is támaszkodni.

Ennek ellentmond a tudományos élet, amely a tudományok térképén nem ismer el hermetikusan záró közigazgatási határokat.

Éppen a mi szaktudományunk is példa erre; tudjuk, hogy mennyire segítette a hangtant mint társadalmi tudományt a természettudományok kategóriájába tartozó anatómia, fiziológia, patológia, akusztika, hogy egyebeket ne említsék. Itt utalok arra, hogy a modern percepciók fonetikában újra felfedezik az (artikulációs) motorikus képzeteknek a beszéd megértésében való szerepét. S ugyancsak modern gondolat (Andreas D. Fröhlich, Hrsg.: Wahrnehmungsstörungen und Wahrnehmungstraining bei Körperbehinderten. Weinstetten, 1977, 9–11), hogy a mozgási fogyatékosággal aránylag többször észlelési zavarok járnak akkor is, ha a mozgászavar nem agyi eredetű; ezek kiterjedhetnek a hallásra (nemcsak a beszédhallásra) és/vagy a látásra. Éppen ezért fel kell figyelnünk arra például, hogy a dadogás, amely mozgási fogyatékoság is, miként hat magának a dadogónak észlelésére általában és a beszédészlelésére.

De hát a nyelvlélektan is hogyan nélkülözhetné a természettudományok segítségét?

## VIII

Ma, a tudományok egyre nagyobb fokú szakosodása idején integrációs tényezők is működnek, s ez biztató jel.

3. Néhány megjegyzés a mai tudomány idolumairól. – Verulami Bacon idolumai közül csak a színház idolumait említem.

A XVII. századi tudós a tudományos tisztánlátást gátló tényezők között említi az "idola theatri"-t; ezeken, a színház idolumain a hagyomány által szentesített, de kritika nélkül elfogadott "tudásunkat" érti.

Ha ma is élne Bacon, "idola theatri"-n alighanem másokat (is) értene. Ezek pedig a tudományos divatok bálványai. A legújabb tudományos divatok bálványainak tiszteletével jár a régebbi szakirodalom tanulmányozásának elhanyagolása, s erre bizony van példa a nyelvtudományban is.

Az erős tudományos egyéniség azonban háttér fordít e bálványoknak, nem enged a pillanatnyi sikerek csábításának, s a tudományos kutatásban is lelkiismeretének categoricus imperativusát követi. Talán remélhetjük, hogy egyre több ilyen tudósunk lesz.

E gondolatok jegyében megnyitom konferenciánkat.

**Vértes O. András**



# A LEXIKÁLIS FONOLÓGIA ELMÉLETÉRŐL

Siptár Péter

MTA Nyelvtudományi Intézete, Budapest

Előadásomban rövid áttekintést adok a Lexikális Fonológia elméletéről, majd bemutatom az elmélet alkalmazását a magyar fonológia egy részterületén, a palatalizációs szabályok kérdéskörén.<sup>1</sup>

Mindenekelőtt szólnom kell néhány szót az elmélet nevééről, mivel a „lexikális” jelző több szempontból is félreérthető. Így például nem valamilyen szélsőségesen konkrét fonológiáról van szó, amely szerint a morféma mindegyik felszíni allomorfja külön-külön fel lenne sorolva a lexikonban; ebből a szempontból a Lexikális Fonológia sztenderd generatív fonológiának számít, bár az is igaz, hogy érdekes megkorlátozásokat tartalmaz az ábrázolások absztraktságával kapcsolatban. Arra is utalhatna a név, hogy az elmélet csak olyan szabályokkal foglalkozik, amelyek a szótári elemek határain belül érvényesülnek. Ez az értelmezés már közelebb áll a valósághoz, de mégsem teljesen igaz, mivel az elmélet egyik állítása éppen az, hogy a fonológia két fő összetevőre oszlik: egy LEXIKÁLIS összetevőre, amelyben a szabályok csak a szó tartományán belül érvényesek, és egy POSZTLEXIKÁLISra, amelyben a szabályok olyan terjedelmesebb szekvenciákon is végbemennek, amelyek a szótári elemeknek a szintaktikai szerkezetbe való beillesztése révén alakulnak ki. A Lexikális Fonológiát tehát azért nevezik így, mert külön tulajdonsághalmazt posztulál a lexikonon belül működő fonológiai szabályok számára. Tekintsük át először a lexikális és posztlexikális szabályalkalmazás közötti különbségeket!

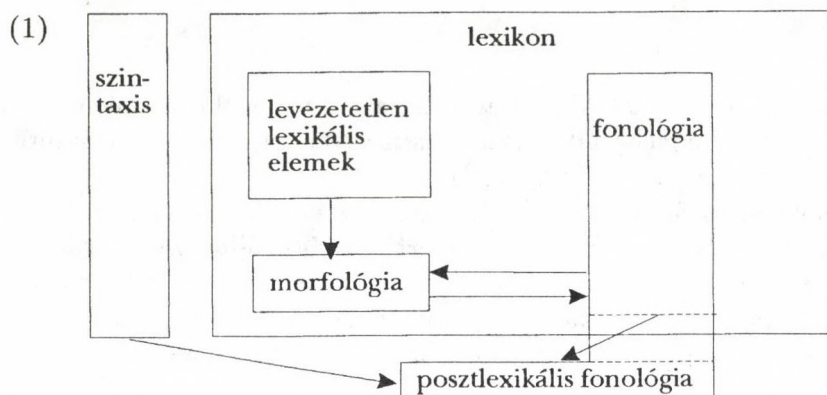
## 1. A posztlexikális szabályok jellegzetességei

(i) Szóhatáron keresztül történő alkalmazás:

A posztlexikális szabályok legkézenfekvőbb ismertetőjele az, hogy nemcsak szavakon belül, hanem szóhatáron keresztül is működésbe lépnek. Az (1) alatti diagramból látható, hogy a szavak nem kapcsolódnak össze monda-

<sup>1</sup> Az elméletet Paul Kiparsky és K.P. Mohanan alkották meg a nyolcvanas évek elején megjelent műveikben (Kiparsky 1982a,b; Mohanan 1982, 1986). Kiváló összefoglalása található Kaisse-Shaw (1985)-ben, amely magyar fordításban is megjelent Siptár (1993) 4. fejezeteként. A nagybetűvel írt „Lexikális Fonológia” az egész elmélet neve, a kisbetűs „lexikális fonológia” a modell egyik komponense, a lexikonon belül működő fonológiai szabályok összessége.

tokká, ennél fogva nem képesek egymás kiejtését befolyásolni, amíg el nem érnek a posztlexikális összetevőbe:



A magyar nyelv egyik tipikus posztlexikális szabálya a zöngésségi hasonulás. A hasonulás szempontjából nem számít, hogy a környezetet alkotó zörejh hang ugyanabban a szóban van-e, mint a szabály által érintett zörejh hang (2a), vagy pedig a következőben (2b):

- (2) a. *tíz*kor [ti:skor]                      b. *tíz* *kör* [ti:skör]

Ezzel szemben egy olyan szabály, amelynek működési tartománya a lexikonon belülre korlátozódik, kizárólag a lexikonon belül létrejövő láncokon fog működésbe lépni. Így például a lexikális *h*-törlés (*cseh*, *csehnek* vs. *csehek*) nem tud arról, hogy a következő szó magánhangzóval kezdődik-e: *cseh ellenzék* [čɛɛl:ɛnze:k], \*[čɛfiɛl:ɛnze:k].

(ii) Kivételnélküliség:

A posztlexikális szabályok alól nem lehetnek lexikális kivételek: ezek mindig működésbe lépnek, ha a szerkezeti leírásuk teljesül.<sup>2</sup> Hasonlítsuk össze az imént említett két szabályt: a *h*-törlés alól számos lexikális kivételt találunk

<sup>2</sup> A Lexikális Fonológia valamennyi változata elfogadja a Zárójeltörlés (Bracketing Erasure) elvét, amelynek az a lényege, hogy a szótári elemek „belügyének” számít mindenféle információ a posztlexikális összetevőbe érve láthatatlanná válik, legyen az a szó morfológiai felépítésére vonatkozó információ vagy az egyes szótári alakokhoz tartozó bármiféle kivételjegyzék.



(*sah, doh, pech, eunuch*), míg a zöngésségi hasonulás alól egyetlen zörejhang-kapcsolat sem vonhatja ki magát, még idegen szavakban, nevekben stb. is könyörtelenül végbemegy (*Macbeth* [gb]).

(iii) Nem-ciklikusság:

Annak az állításnak, hogy a posztlexikális szabályok nem tudnak hozzáférni a szavak belső felépítéséhez, az előbbiektől független bizonyítéka (folyománya), hogy nem alkalmazhatók ciklikusan (a ciklikus szabályalkalmazásról vö. Siptár 1985, 1988-t és az ott idézett irodalmat). A ciklikus alkalmazás lehetősége természetesen attól függ, hogy vannak-e egymásba ágyazott tartományok, ámde a szavak a belső tartományokra utaló nyomokat teljesen elveszítik, mire előbújnak a lexikális összetevőből:

- (3) a. lexikális ábrázolás (bemenet): [tana]tVk  
 b. posztlexikális ábrázolás: [tanítsatok]

(iv) Új alakzatok létrehozása:

A posztlexikális szabályok létrehozhatnak újfajta szekvenciákat, szerkezeteket vagy szegmentumokat, tehát pl. olyan szegmentumokat, amelyek nincsenek benne a mögöttes szegmentumok tárában, vagy olyan tulajdonságaik vannak, amelyek a mögöttes ábrázolásban nem állnak rendelkezésre. (A lexikális szabályok ezzel szemben SZERKEZETÖRZŐK.) A szóvégi *j* zöngétlenedése (*kapj, rakj, döfj*) például [ç]-t hoz létre, amely a mögöttes mássalhangzó-rendszerben nem szerepel. Az Ikeredés szabálya, amely a geminátákból felszíni hosszú mássalhangzót hoz létre, bevezeti a [hosszú] jegyet, amely a mögöttes mássalhangzókészlet jellemzésében nem vesz részt; ugyanígy a nazálisok képzési hely szerinti hasonulása bevezeti az [osztott] (distributed) jegyet. A posztlexikális szabályok olyan szekvenciákat is létrehozhatnak, amelyek előfordulása a lexikonon belül nem megengedett. Így például az az angol gyorsbeszéd-szabály, amely törli a svát a szókezdő szótagokból, teljesen szokatlan szókezdő mássalhangzó-kapcsolatokat hozhat létre (*potato* [pteitou] 'krumpli'). A posztlexikális szabályok ezen kívül fokozatos (nem-bináris) kimenetet is létrehozhatnak; továbbá csak ezek lehetnek fakultatívak és függhetnek a beszédsebességtől.

(v) Alkalmazásuk nem-levezetett alakokra:

A posztlexikális szabályokra nem érvényes a Szigorú Ciklus Elve. A SzCE-t a későbbiekben részletesebben is tárgyalni fogjuk; egyelőre annyi is elegendő, hogy a SzCE egyik hatása abban áll, hogy megakadályozza a lexikális szabályok alkalmazását olyan szekvenciákra, amelyek már mögöttesen megfelelnek a szerkezeti leírásuknak (például a /h/ nem törlődik az *ihlet, jacht*,

*technika* típusú szavakban). Ez a megszorítás a posztlexikális összetevőben már nem érvényesül: így például azok a jövevényszavak, amelyek zöngésség tekintetében eltérő zörejhang-kapcsolatot tartalmaznak, hasonuláson esnek át: *anekdota* [gd], *afgán* [vg], *abszolút* [ps], *joghurt* [kh], stb.

(vi) A lexikális ábrázolás:

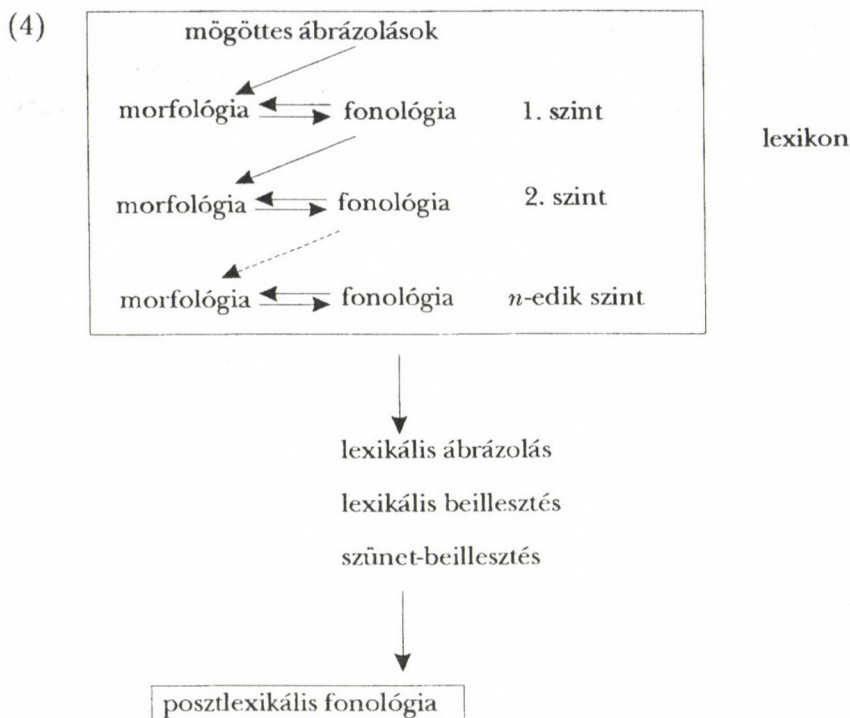
Mohanán (1982) hangsúlyozza, hogy a fonológia kettéosztása lexikális és posztlexikális részre a levezetés egy jelentős szintjének létezését jósolja meg. Ez a szint a lexikon kimenete, amelynek több érdekes tulajdonsága van. Ezek egy része ennek a közbülső ábrázolásnak (amelyet Mohanán LEXIKÁLIS ÁBRÁZOLÁSNAK nevez<sup>3</sup>) a pszichológiai realitásával kapcsolatos. Így például a beszélők a lexikális szabályok kimenetét különösebb előképzettség híján is különbözőnek ítélik a bemenetüktől, tehát a magyar beszélők könnyen meg tudják különböztetni a *lássa* és a *látja* alakokat; másfelől azonban a /h/ valamennyi allofónját, a zöngés [h] -t és a veláris [x] -t is, a zöngétlen glottális [h] -val azonosnak vélik. Mohanán azt is állítja, hogy a nyelvbtlásokra a lexikális szabályok alkalmazását követően kerül sor, következésképp nem kell velük összhangban állniuk; a posztlexikális szabályok viszont a nyelvbtlások alakokon is végbemennek. Végül a lexikális szabályokat nem hiúsítják meg a szünetek, a posztlexikálisakat viszont igen. Így a *háttal* alak közbeékeltszünettel is *hát...tal*-nak hangzik, míg a *hátba* nem *hád...ba*, hanem *hát...ba* lesz. Mohanán ezt a jelenséget úgy magyarázza, hogy a szünet beillesztését az első posztlexikális szabálynak tekinti. Így a -*val* első mássalhangzója még szomszédos a *hát* végével a hasonulás alkalmazásakor, de a másik esetben a *hát* végén álló mássalhangzó már nem áll zöngés zörejhang előtt, mire a zöngésségi hasonulásra kerülne a sor.

## 2. Morfológiai és fonológiai rétegek a lexikonon belül

A lexikális és posztlexikális összetevők közötti viszonylatok legtöbbje kicsiben megismétlődik a lexikonon belül felállított rétegek közötti viszonylatok formájában, amint (4) mutatja.

<sup>3</sup> A „lexikális ábrázolás” kifejezés a sztenderd generatív fonológiában a mögöttes ábrázolásra szokott utalni; itt azonban azt az alakot kell érteni rajta, amelyen az összes alkalmazható morfológiai és lexikális fonológiai szabály végbement már, de posztlexikális szabály még egy sem.





Így például mindegyik réteg kimenete egy-egy (levezetett) szótári egység, akár csak a teljes lexikon kimenete. A zárójeltörlés, amelyre az imént utaltam, az egyes rétegek között is érvényes:

(5) Zárójeltörlési konvenció

A belső zárójelek mindegyik szint végén törlődnek

(Kiparsky 1982a,140).

Más szóval: az egy-egy réteghez tartozó információ hozzáférhetetlen a következő réteg számára. A rétegek pontos száma részint nyelvspecifikus, részint egy-egy nyelven belül is vitatott. Például az angolban Kiparsky (1982a,b) szerint kettő, Halle és Mohanan (1985) szerint négy réteget kell megkülönböztetnünk. Amint az előadás második felében a palatalizációs szabályokkal kapcsolatban látni fogjuk, a magyarban legalább két szint van – de vajon van-e szükség kettőnél többre? A szóösszetételek Halle és Mohanan modelljében (az angolban) a 3. szinten alakulnak ki. És a magyarban? Amennyiben feltesszük, hogy a *h*-törlés a magyarban 1. szintű szabály, a szóösszetételeknek legkorábban a 2. szinten szabad kialakulniuk (vö. pl. *Cseh-*

*ország* \*[čɛɦɔrsa:g], *Cseh-erdő* \*[čɛɦɛrdö:]). Nincs olyan bizonyíték, ami arra utalna, hogy a *h*-törlés nem lehet 1. szintű szabály. A Lexikális Palatalizáció szabályáról azonban látni fogjuk, hogy 1. és 2. szintű szabály. Mivel azonban összetétel-határon keresztül nem működik (vö. pl. *látjátok* [la:tʲ:a:tok], de *átjáró* \*[a:tʲ:a:ro:]), ez arra utal, hogy a szóösszetételek kialakulása a 3. szintre tehető.

### 3. Palatalizáció

Palatalizációnak nevezzük azt a fonológiai folyamatot, amelynek során egy palatális mássalhangzónak, tehát a /j tʲ dʲ nʲ/ mássalhangzók egyikének hatására az előtte álló (dentális) mássalhangzó (palato-alveoláris vagy) palatálissá válik.<sup>4</sup>

A magyar fonológiában három különböző mélységi szinten találunk palatalizációs szabályokat. Ezek: a *t*-végű igék felszólító alakjaiban található morfofonológiai palatalizáció (*üt/üss, önt/önts*), a *látja, hordja, fonja, falja*-típusú lexikális palatalizáció, valamint az a (fakultatív) posztlexikális palatalizáció, amely az *átjáró, védjegy, van joga, feljön*-féle példákban figyelhető meg. Ez utóbbi (a fakultativitástól eltekintve) abban különbözik a másik kettőtől, hogy ezúttal a kimenet nem olvad össze a szabály alkalmazását kiváltó szegmentummal (vö. *menetjegy* [mɛnɛtʲɛdʲ]~[mɛnɛtʲɛdʲ], \*[mɛnɛtʲ:ɛdʲ]), és hogy a kiváltó nemcsak /j/, hanem palatális zárhang vagy nazális is lehet, pl. *van gyufa* [nʲdʲ], *két nyúl* [tʲnʲ] (a részleteket l. Siptár 1990-ben).

Kezdjük a *t*-végű igék felszólító módú alakjával. A felszólító mód *j*-je előtt a tövégi *t*-vel háromféle dolog történhet:

- [š]-ként jelenik meg rövid magánhangzó után, pl. *üt/üss*;
- [č]-ként jelenik meg, ha előtte zengőhang áll, pl. *hajt/hajts, önt/önts, olt/olts*;
- törlődik, ha előtte zörejhang áll. Ez a zörejhang kétféle lehet: /š/ a *fest* szóban, illetve /s/ az *oszt*-típusú szavakban (Abondolo 1988:146 szerint az ÉrtSz 117 ilyen igetövet tartalmaz, valamennyiben morfémahatár található az *sz* és a *t* között).

A hosszú magánhangzó + *t* végű igék közül a *lát*, a *bocsát* és a *lót* (-*fut*) a rövid magánhangzósokhoz hasonlóan, a *fűt, hűt, műt, szít, tát*

<sup>4</sup> Tehát nem értjük bele azt a teljesen automatikus, alacsonyszintű, nem-neutralizáló (és feltehetőleg nem is nyelvspecifikus) típusát a „fonetikai palatalizációnak”, amelyet a nem-alsó nyelvtanulási előképzett magánhangzók és a /j/ váltanak ki, és amelynek révén többékevésbé palatalizált velárisok/dentálisok/labialisok jönnek létre, mint például a *kín* szó *k*-ja (szemben a *kár* szó sima és a *kút* szó kerekített *k*-jával). Az efféle eseteket az alábbiakban nem tekintjük palatalizációnak.



és *vét*, valamint az *-ít* képzős igék pedig a zengőhang + *t* végűekhez hasonlóan viselkednek (vö. Jakab 1967). Ezt a megoszlást elvben háromféleképpen írhatjuk le. Először is, megelégedhetünk az egyszerű felsorolással, ahogyan az imént tettük. Ez a megoldás a legegyszerűbb, de a legkevésbé vonzó. Másodszor: azt mondhatjuk, hogy a *lát*-félék kivételesek, és a hosszú magánhangzós igetövek szabályosan [č]-re változtatják a tövégi *t*-t: ez a MMNyR (Tompá 1961, I,490) s hasonlóképpen Abondolo (1988,146) megoldása. Végül Vago (1980,72) nyomán feltehetjük, hogy a *lát*-félék a rövid magánhangzósokkal együtt alkotják a szabályos (magánhangzó + *t*) esetet, a *fűt*-féléket pedig valahogyan bele kell illeszteni a [č]-osztályba. Vago ezt úgy éri el, hogy az érintett igék mögöttes ábrázolásába felvesz egy /j/-t, amely a *t* (most már) szabályos [č]-vé palatalizálódása után – de még a *hajts*-félékben bekövetkező degemináció előtt – a megelőző magánhangzóba beleolvad (illetve annak pótlónyúlása mellett törlődik): /fűjt+j/ → fűjč: → [fű:č:] *fűts*. Ez a megoldás nem annyira önkényes, mint amilyennek első pillantásra tűnik: a felsorolt igék számos más szempontból is úgy viselkednek, *mintha* /Vjt/- végűek lennének. Például a főnévi igenév végződése magánhangzó + *t* végű igékhez kötőhang nélkül kapcsolódik: *űtni*, *futni*, *lítani*, mássalhangzó + *t* végűekhez viszont kötőhanggal: *osztani*, *hajtani*, *váltani*. A *fűt*-félék az utóbbi csoporthoz csatlakoznak: *fűteni*, *hűteni*, *szítani*, stb.<sup>5</sup>

Az egyetlen pont, ahol némi önkényesség van, a feltételezett mögöttes zengőhang minősége: miért éppen /j/? Tulajdonképpen kizárásos alapon: ez az egyetlen szóhajózható zengőhang, amely a felszínen nem jelenik meg ilyen környezetben (vö. Vago 1991). Az autoszegmentális, illetve CV-fonológia keretében ez a kevéske önkényesség is eltávolítható (vö. Vago 1987, 1989, 1991):

$$(6) \quad \begin{array}{ccccc} \text{tát} & = & \text{CVCC} & & \text{lát} = \text{CVVC} \\ & & | \vee | & & | \vee | \\ & & \text{t} \text{ a } \text{t} & & \text{l} \text{ a } \text{t} \end{array}$$

Összefoglalva tehát az eddigieket (formalizálva l. Vago 1980:71):

$$(7) \quad \begin{array}{ll} \text{a. } t \longrightarrow \check{s} / V \_\_ j \\ \text{b. } t \longrightarrow \check{c} / \left[ \begin{array}{c} C \\ + \text{szon} \end{array} \right] \_\_ j \\ \text{c. } t \longrightarrow \emptyset / [-\text{szon}] \_\_ j \end{array}$$

<sup>5</sup> Megjegyzendő azonban, hogy a *bocsát* igét a *-ni*-teszt a *fűt*-félékhez sorolja, ráadásul *bocsájt* alakváltozata is létezik, a felszólító alakja mégis [š]-t, nem [č]-t tartalmaz.

Most a következő két kérdés áll előttünk:

1. Mi történik a /j/-vel?

2. Hogyan korlátozható mindez a felszólító módra?

A felszólító mód /j/-je a fenti palatalizációs folyamatok révén [š], [č], illetve [s] utáni helyzetbe kerül, és ugyanaz történik vele, mint a szibiláns-végű igék felszólító módú alakjában: *mos+j* → *moss*, *fut+j* → *fus+j* → *fuss*. Más szóval a *t*-palatalizáció/törlés táplálja a *j*-hasonulás szabályát; a következő levezetések ezt szemléltetik:

(8)	/moš+j/	/fut+j/	/hajt+j/	/fešt+j/
<i>t</i> -pal/törl. (7)	–	š	č	∅
<i>j</i> -hasonulás (9)	š	š	č	š
degemináció	–	–	∅	–
egyéb szabályok	[moš:]	[fuš:]	[hojč]	[feš:]

A *j*-hasonulás Vago (1980:36) alapján így fogalmazható meg:

$$(9) \quad \begin{matrix} [+ \text{ érdes}] & j \\ 1 & 2 \end{matrix} \longrightarrow \begin{matrix} 1 & 1 \end{matrix}$$

A *t*-palatalizáció/törlés a sztenderd magyar fonológiában csak felszólító igealakokban megy végbe.<sup>6</sup> Ezt a kikötést bele lehet írni magába a szabályba, de szerencsésebb lenne valamilyen elvi indoklást találni hozzá. Vago (1980) megoldása szerint az igei és birtokos személyragok *j*-je egyaránt betoldott, nem mögöttes szegmentum (más-más szabály toldja be őket, Vago 1980,68; 105); ha a betoldó szabályok a *t*-palatalizáció/törlés után kerülnek alkalmazásra (akár extrinszikus rendezés révén, akár valamilyen más okból), a *t*-pal/törl. korlátozása a felszólító alakokra automatikusan megoldódik.<sup>7</sup> A Lexikális Fonológia keretében kínálkozik egy másik megoldás, amely sem betoldó szabályokat, sem extrinszikus rendezést nem igényel: eszerint a felszólító mód jele az első szintű (mélyebb) toldalékok egyike, a személyragok a második (felszínhez közelebb) szintre tartoznak (ezt összefűzésük sorrendje

<sup>6</sup> A *suk-sükkölés* néven ismert sztenderd-alatti változatban a *j*-t tartalmazó igei személyragok előtt, tehát kijelentő módú alakokban is lejátszódik; a birtokos személyragok *j*-je azonban egyetlen magyar nyelvváltozatban sem váltja ki a szabály alkalmazását: *botja* \*[boš:ɔ], *lantja* \*[lɔnčɔ].

<sup>7</sup> A birtokos *j* minden bizonnyal valóban betoldott, vö. Kiefer (1985); és lehet, hogy az igei személyragok *j*-je is annak bizonyul – a jelen gondolatmenet elsősorban nem a 'betoldott/nem betoldott' dilemmáról szól.



is tanúsítja). A *t*-palatalizáció/törlés szabálya csak az első szinten működik, tehát a személyragok *j*-je nem képes működésbe lendíteni:

(10)	(felsz.)	(kij.)	(kij.)
	/la:t+j+juk/	/moš+juk/	/la:t+juk/
1. szint	la:t+j	–	–
<i>t</i> -p/ <i>t</i>	š		
<i>j</i> -has	š		
2. szint	la:šš+juk	moš+juk	la:t+juk
<i>j</i> -has	š	š	
pal			t <sup>y</sup> t <sup>y</sup>
deg	∅		
	[la:š:uk]	[moš:uk]	[la:t <sup>y</sup> :uk]

Jellegzetes vonása a fenti megoldásnak, hogy a *j*-hasonulásnak mindkét szinten működésbe kell lépnie. Miért nem lehetséges az, hogy csak a 2. szinten kezdjen működni? Ezt kizárja a Szigorú Ciklus Elve, amelyet a következőképpen fogalmazhatunk meg (Siptár 1988,5):

- (11) Egy R ciklikus szabály akkor működik helyesen *j* ciklusban, ha vagy (a) vagy (b) teljesül:

- (a) R olyan információt használ fel, amely részben hozzáférhető volt a ciklikus szabályok egy korábbi végigfuttatása során, részben pedig először a *j* ciklusban válik hozzáférhetővé.

Az (a) feltétel három különböző esetet foglal magában:

R kimondottan utal valamely A-ra és B-re, az alábbiakban:

- (i) [<sub>j</sub> XAY...[<sub>j-1</sub>...B...] Z];  
(ii) [<sub>j</sub> [<sub>j-1</sub>...B...]...XAY];  
(iii) [<sub>j</sub> X [<sub>j-1</sub>...A...] Y [<sub>j-1</sub>...B...] Z].

- (b) R olyan információt használ fel, amelyet a *j* ciklusban egy R előtt alkalmazott szabály hoz létre.

A ciklikus szabályok tehát nem működhetnek levezetetlen alakokban; például a *Vászja* név nem hangzik *Vássszá*-nak. Továbbá, mivel a belső zárójel minden szint végén törlődnek, a következő szint szabályai nem tudnak különbséget tenni egy teljesen levezetetlen alak (mint *Vászja*) és egy előző szinten létrejött alak (mint a (10)-beli *lás-j*) között. Így azután – a (11b) szerinti, itt nem fennálló esettől eltekintve – nem tudnak benyúlítani az előző

szinten már lezárult területre, csak az éppen szóban forgó morfémahatáron tudnak tevékenykedni. Ezért kell a *j*-hasonulás szabályának a felszólító mód jelére már az 1. szinten működnie; a *-juk* személyrag viszont csak a 2. szinten kapcsolódik az alakhoz, tehát a *j*-hasonulás innen sem hiányozhat. A Lexikális Fonológia elvei szerint azonban semmi akadály, hogy egy-egy szabály többszintű legyen – egynél több szinten is megjelenjen –, de csak akkor, ha az érintett szintek szomszédosak. Ez a feltétel a jelen esetben teljesül.

Térjünk át a palatalizáció következő típusára, amelyet a (10) alatti levezetésben már felhasználtunk. Ezt a fajta palatalizációt (pl. *látja*, *adja*, *kenje*, *falja*) ugyancsak a /j/ váltja ki, de ezúttal az eredmény valóban palatális hang, mégpedig gemináta; a szabály bemenetét a dentálisok (/t d n l/) és – részben üresjáratban – a palatálisok (/tʲ dʲ nʲ/) alkotják. A folyamatot leírhatjuk szabáлыпárral, a fenti *t*-pal/törl. és *j*-hasonulás mintájára, vagy egyszerre, összeolvadásként. Nézzük előbb a sztenderd megoldást:

(12) Palatalizáció

$$\begin{bmatrix} + \text{kor} \\ - \text{érd} \end{bmatrix} \longrightarrow [- \text{elő}] / \_\_\_ j$$

(azaz: *j* előtt *t*-ből *ty*, *d*-ből *gy*, *n*-ből *ny*, *l*-ből *j* lesz)

(13) Palatális-összeolvadás (PÖ)

$$\begin{bmatrix} - \text{elő} \\ + \text{kor} \\ - \text{érd} \end{bmatrix} \quad j \quad \begin{matrix} 1 & 2 \end{matrix} \longrightarrow \begin{matrix} 1 & 1 \end{matrix}$$

(azaz: *tyj*-ből *tty*, *gyj*-ből *ggy*, *nyj*-ből *nyy* lesz)

E megoldás előnye, hogy az első szabály némi átformálásával belevonható a posztlexikális palatalizáció is (pl. *mit jelent*, *családjog*, *színjeles*, *feljön*; *két nyúl*, *van gyufa*, *vad tyúk*), amely bármely palatális mássalhangzó (/j tʲ dʲ nʲ/) előtt bekövetkezik, de nem jár összeolvadással (csak ikeredéssel, ha a kimenet véletlenül két azonos mássalhangzót tartalmaz: *hét tyúk*, *vad gyűlölet*, *vén nyárfa*). Így posztlexikális esetekben csak a (megfelelően általánosított) első szabály, lexikális esetekben mindkét szabály alkalmazásra kerülne; más szóval, a PÖ a lexikonon belülre lenne korlátozva, a Palatalizáció viszont itt is, ott is alkalmazható volna. További támogatással szolgál ehhez a megoldáshoz az a néhány (igaz, marginális) eset, amikor szón belül van szükség a palatalizáló szegmentumoknak a /j/-nél tágabb körére: *pattan* – *pattantyú* [nʲtʲ], *pillanat* – *pillanatnyi* [tʲnʲ].



Ekkor tehát a Palatalizáció szabálya így fest (vö. Vago 1980,40):

$$(14) \begin{bmatrix} + \text{ kor} \\ - \text{ érd} \end{bmatrix} \longrightarrow [- \text{ elöl}] / \text{ — } \begin{bmatrix} - \text{ elöl} \\ + \text{ kor} \\ - \text{ érd} \end{bmatrix}$$

(azaz: *ty, gy, ny, j* előtt  $t \rightarrow ty, d \rightarrow gy, n \rightarrow ny, l \rightarrow j^8$ )

Ebben a formában a szabály most már korlátozás nélkül alkalmazható a lexikális és posztlexikális összetevőben egyaránt; ezen belül lexikálisan az első és a második szinten kötelezően, a posztlexikális összetevőben pedig fakultatívan. A (10) alatti levezetések által sugalltakkal ellentétben az első szintről sem zárhatjuk ki (gondoljunk az SzCE-re), hiszen a felszólító mód /j/-je is kiváltja, ha a tövégi mássalhangzó /d n l/: *adjon, kenjen, keljen*; a tövégi /t/-k az egyébként-elv (más néven: a teljes bennfoglalás elve) révén automatikusan a Palatalizáció elé rendezett *t*-pal/törlésen már átesve kikerültek a szabály lehetséges bemenéteinek köréből. Tovább szépíthető ez a megoldás a (9) *j*-hasonulás és a (13) PÖ összevonásával (így Abondolo 1988,64 is):

(15) *j*-hasonulás (általánosítva)

$$\begin{bmatrix} + \text{ kor} \\ \left\{ \begin{array}{l} - \text{ elöl} \\ + \text{ érd} \end{array} \right\} \end{bmatrix} \quad j \quad \begin{matrix} 1 & 2 \end{matrix} \longrightarrow \begin{matrix} 1 & 1 \end{matrix}$$

(azaz: *szj* → *ssz*, *zj* → *zz*, *cj* → *cc*, *dzj* → *ddz*, *sj* → *ss*,  
*zsj* → *zzs*, *csj* → *ccs*, *dzsj* → *ddzs*, *tyj* → *tty*,  
*gyj* → *ggy*, *nyj* → *nny*)

Az így összevont *j*-hasonulás tehát első + második szintű lexikális szabály, a Palatalizáció pedig minden szinten (és posztlexikálisan is) alkalmazható.

<sup>8</sup> Az *l* posztlexikális palatalizációja (alkalmazásának szintaktikai feltételrendszeréről vö. Vogel–Kenesei 1988) korlátozottabb, mint a többi bemeneté: csak /j/ előtt következik be (*feljön* [lj]~[j:], de *felnyal* \*[jnʸ], *ökölnyi* \*[jnʸ], *kopoltyú* \*[jtʸ]). Olsson (1992) ezért azt javasolja, hogy az /l/ legyen teljesen kizárva a palatalizációból, és a rákövetkező /j/-hez való teljes hasonulásáról (a lexikális-kötelező *éljen* és a posztlexikális-fakultatív *hol jártál* esetében egyaránt) egy, a palatalizációtól független szabály adjon számot. Megjegyzendő továbbá, hogy a palatalizáció posztlexikális alkalmazásait különféle egyéb posztlexikális szabályok gátolhatják (bleed); ilyen az /l/-törlés (a megelőző magánhangzó pótlónyúlásával vagy – a beszédtempótól függően – anélkül, pl. *eljön* [e:jön]~[ejön]), valamint az /n/-vokalizáció szabálya, amely folyamatos mássalhangzók (a réshangok és /l r j h/) előtt törli az /n/-t, miközben a megelőző magánhangzót nazalizálja, pl. *Bán Jani* [bã:jɔni] (n-vokalizáció) vs. *bã:nja* [ba:nʸ:ɔ] (lexikális palatalizáció).

Befejezésül vizsgáljuk meg röviden a másik lehetőséget: a (12) és (13) összevonását egyetlen (összeolvadási) szabállyá; ekkor a *j*-hasonulás maradna (9) szerint, külön posztlexikális palatalizációs szabály is kell kb. (14) szerint; a *pillanatnyi*-féle esetek minden további nélkül a posztlexikális komponensbe utalhatók, ahol – esetleg a *rongy*, *satnya* típusú morfémán belüli esetekkel együtt – végbemegy rajtuk a palatalizáció.

Maga az összeolvadásként felfogott (lexikális) palatalizációs szabály így festene:

$$(16) \begin{bmatrix} + \text{ kor} \\ - \text{ érd} \end{bmatrix} \quad j \quad \begin{matrix} 1 & 2 \end{matrix} \longrightarrow \begin{bmatrix} 1 \\ - \text{ elöl} \end{bmatrix} \quad \begin{bmatrix} 1 \\ - \text{ elöl} \end{bmatrix}$$

(azaz:  $tj \rightarrow tty$ ,  $dj \rightarrow ggy$ ,  $nj \rightarrow nny$ ,  $lj \rightarrow jj$ ;  
 $tyj \rightarrow tty$ ,  $gyj \rightarrow ggy$ ,  $n\ddot{y}j \rightarrow nny$ )

Eszerint tehát az érintett változások egy lépcsőben, a posztlexikális palatalizációtól teljesen függetlenül valósulnak meg. A kétféle megoldás közötti választás jórészt elméletfüggő; bár mindkettő megfogalmazható klasszikus generatív és autoszegmentális keretben egyaránt (az utóbbira vö. Siptár 1992), az előbbiben az első, az utóbbiban a második megoldás tűnik kézenfekvőbbnek. A Lexikális Fonológia szempontjából azonban mindkét megoldás egyenértékű, és megfordítva: a magyar palatalizáció itt megfigyelt jelenségei bármelyik megoldás esetén tökéletesen illeszkednek a Lexikális Fonológia elméletének predikcióihoz.

## Irodalom

- Abondolo, D.: Hungarian Inflectional Morphology. Budapest 1988.
- Halle, M. – Mohanan, K. P.: The segmental phonology of Modern English. *Linguistic Inquiry* 16. 1985. 57–116.
- Jakab László: A *t* végű igék felszólító módjához. *Magyar Nyelv* 63. 1967. 194–196.
- Kaisse, E. M. – Shaw, P.: On the theory of Lexical Phonology. *Phonology Yearbook* 2. 1985. 1–30.
- Kiefer F.: The possessive in Hungarian: A problem for Natural Morphology. *Acta Linguistica* ASH 35. 1985. 85–116.
- Kiparsky, P.: From cyclic phonology to lexical phonology. In: H. van der Hulst–N. Smith szerk., *The structure of phonological representations*. Dordrecht I. 1982a. 131–175.
- Kiparsky, P.: Lexical morphology and phonology. In: I.S. Yang szerk., *Linguistics in the Morning Calm*. Seoul 1982b. 3–91.

- Mohanan, K. P.: Lexical phonology. MIT doktori disszertáció, terjeszti az Indiana University Linguistics Club. 1982.
- Mohanan, K. P.: The theory of lexical phonology. Dordrecht 1986.
- Olsson, M.: Hungarian Phonology and Morphology. Lund 1992.
- Siptár P.: Palatalization in English and the Strict Cyclicity Principle. In: J.C. Bognár–A.C. Rouse szerk., Proceedings of the Second Conference on Speech Orientated ELT. Pécs 1985. 222–235.
- Siptár P.: Cyclicity and palatalization. Magyar Fonetikai Füzetek/Hungarian Papers in Phonetics 19. 1988. 5–13.
- Siptár Péter: Palatalizáció. Részlet egy készülő magyar fonológiából. In: Földi É. szerk., Studia in Honorem K. Bolla Oblata a Collegis et Discipulis. Egyetemi Fonetikai Füzetek 3. 1990. 149–158.
- Siptár Péter: A magyar mássalhangzórendszer autoszegmentális megközelítése: ábrázolások és szabályok. Nyelvtudományi Közlemények 93. 1992. (megj. előtt)
- Siptár Péter szerk.: Modern fonológiai szöveggyűjtemény II. Szeged 1993.
- Tompa József szerk.: A mai magyar nyelv rendszere. Budapest 1961.
- Vago, R. M.: The sound pattern of Hungarian. Washington 1980.
- Vago, R. M.: On the representation of length. In: W. Dressler et al. szerk., Phonologica 1984. Cambridge 1987. 319–324.
- Vago, R. M.: Compensatory lengthening in Hungarian: Implications for mora theory. Kézirat. 1989.
- Vago, R. M.: Elméleti szempontok a *t* végű igék elemzésében. In: Kiss J.–Szűts L. szerk., Tanulmányok a magyar nyelvtudomány történetének témaköréből. Budapest 1991. 682–690.
- Vogel, I. – Kenesei, I.: The interface between phonology and other components of grammar: the case of Hungarian. Phonology Yearbook 4. 1988. 243–263.



# A LEXIKÁLIS HOZZÁFÉRÉS

Gósy Mária

MTA Nyelvtudományi Intézete, Budapest

## Bevezetés

A beszédmegértés során akusztikai jelsorozatok feldolgozása történik meg, amelyben a hallgató mindazon nyelvi ismeretét (fonémákat, jelentést, szintaktikai szerkezeteket stb.) felhasználja, amelyek segítségével a megértéshez eljut. Az alapkérdés az, hogy miként jön létre a megfeleltetés a szót reprezentáló akusztikai jelsorozat és az egyén mentális lexikona között, illetőleg az akusztikai jelsorozat és a mentális lexikon egy adott egysége között. Az a folyamat, amelynek eredményeként ez megtörténik, a lexikális hozzáférés, azaz a szóértés. Az angolszász szakirodalomban a mentális lexikon az egyéni szókincset, annak aktív és passzív részét jelenti. Megítélésem szerint egy harmadik dimenziót is értelmeznünk kell, amikor ezt a kifejezést használjuk: az aktivizálható szókincs fogalmát, amely lényegét tekintve mind az aktív, mind a passzív szókincstől különbözik (ez utóbbi kettő nem szorul magyarázatra). Az aktivizálható szókincs a következőképpen definiálható: azon szavak összessége, amelyek a beszédpercepció folyamatokban a kiváltó (akusztikai) inger hatására extra energiával töltődnek fel, azaz aktivizálódnak. Az aktivizálódás érinthet sok, több vagy egyetlen szót a mindenkori akusztikus input és az átviteli közeg paramétereinek függvényében.

Tekintettel arra, hogy a lexikális hozzáférés folyamata közvetlenül nem, illetőleg csak részlegesen vizsgálható; elkerülhetetlen, hogy modellek felállításával próbáljuk a működési folyamatokat megismerni. Az ismert modellek különféle elméletek reprezentánsai, ennek megfelelően kisebb-nagyobb mértékben különböznek egymástól. Többségük nyelvfüggetlen, azaz a mechanizmusnak az általános emberi vonatkozásait hangsúlyozzák. Egyre nyilvánvalóbb azonban, hogy a sajátosan nyelvi elemzése és a modellbe építése nélkül a lexikális hozzáférési folyamatok csak egy bizonyos pontig ismerhetők meg. Az újabb kutatások éppen ezért fordulnak az általános mechanizmusba épített nyelvspecifikus folyamatok megismerése felé.

## A folyamat működése

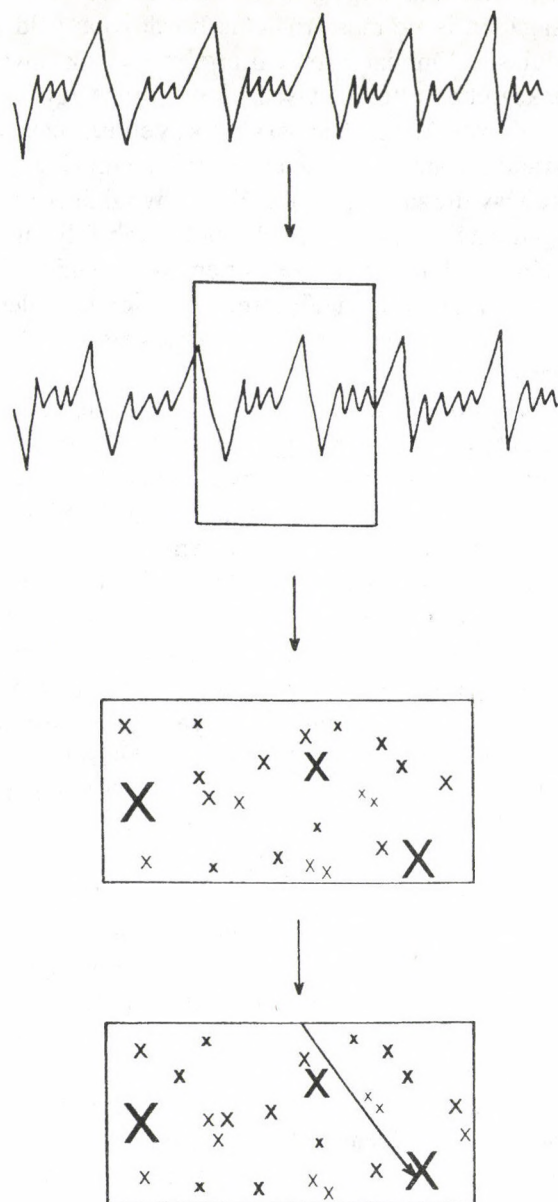
A lexikális hozzáférés folyamatában biztosan ismertek a bemenet (input) és a kimenet (output) jellegzetességei, legalábbis amennyire ezek objektív értékekkel meghatározhatók. Az input maga az akusztikai jelsorozat, amely elemezhető és az idő, a frekvencia és az intenzitás függvényében definiálható. Az elemzési eredmények azonban nemcsak a (felismeréshez) szükséges, tehát az invariáns, ha-



nem a redundáns, sőt az irreleváns paramétereket is tartalmazzák, amelyeknek az egymástól való elkülönítése és eltérő felhasználása éppen abban a folyamatban történik, amely vizsgálódásunk tárgya, a lexikális hozzáférés során. Az input adattömeg már önmagában is számos probléma hordozója; például az egyes beszédhangok a dekódolás különböző szintjein értelmezhetők; invariáns jegyeik a kontextushatás következtében eltérően viselkednek; nincs egy-az-egynek megfelelés az artikulációs változás és annak akusztikai következménye tekintetében; ugyanazon szó akusztikai paraméterei jelentősen módosulnak az idő (beszédtempó), az artikuláció (ejtésváltozatok), az aktuális szöveggörnyezet és más tényezők (pl. emóció) következtében (s ez a felsorolás tovább folytatható). A lexikális hozzáférés folyamatának állandóan készenlétben kell lennie mindezen problémák fogadására és megoldására egyfelől a mentális lexikon eléréséig, másfelől pedig a lexikonban való műveletvégzésig, amelynek végső eredménye az output, avagy a szófelismerés (vö. 1. ábra).

A szófelismerés folyamatának vizsgálata kezdetekben az izolált szavakra koncentrált. Ezzel jelentős mértékben leegyszerűsödött a probléma, hiszen a folyamatos beszéd következményeként létrejövő akusztikai-fonetikai jelenségek nem érvényesültek (koartikuláció, ill. kontextushatás). Ráadásul a keresett szavak többnyire megfeleltek a szótári formáknak, ennek a predikciónak a sugallásával pedig a természetestől eltérően működő lexikális hozzáférést vizsgáltak, illetőleg a folyamat korlátozott – bár egyfajta lehetséges – működéséről nyertek információkat. Az ún. szótári alakok és a folyamatos beszédben előforduló szavak között az agglutináló nyelvek esetében lényegesen nagyobb a különbség a nem agglutinálókhöz képest, s ez döntően befolyásolja a felismerési folyamatokat. Az izolált szavak felismerésére vonatkozó kutatások ugyanakkor fontos megállapításokat tettek, csak azok érvényessége maradt korlátozott. Kísérleti adatok igazolták hol az akusztikai szerkezet, hol a szógyakoriság prioritását a szófelismerésben, míg mások a szöveggörnyezet elsőségét hangsúlyozták (Nooteboom – Truin 1980; Tanenhaus – Lucas 1987; Forster 1989; stb.).

A lexikális hozzáférés vizsgálatát a folyamatos beszédben rendkívüli módon megnehezíti az a tény, hogy a dekódolási folyamat valamennyi szintjének működésével kell számolnunk, s ezek izolálása rendkívül nehéz (hacsak nem lehetetlen) feladat. Továbbá az első nehézség már a kezdeteknél jelentkezik: megoldatlan a szegmentálás problémája. Az artikuláció folyamatossága következtében létrejövő folyamatos akusztikai jelsorozat ugyanis nem tartalmaz fogódzókat arra vonatkozólag, hogy hol kezdődnek, illetve hol végződnek az egyes szemantikai egységek. Így tehát a szegmentálás is a lexikális hozzáférés egyik funkciója kell, hogy legyen, amely pillanatnyi tudásunk szerint csak kognitív szinten oldható meg. Az izolált szavak felismerésekor nincs szegmentálási nehézség, hiszen az akusztikai jelsorozat kezdete és vége egyúttal a szemantikai egység kezdete és vége is.



1. ábra

A lexikális hozzáférés folyamatának sémája

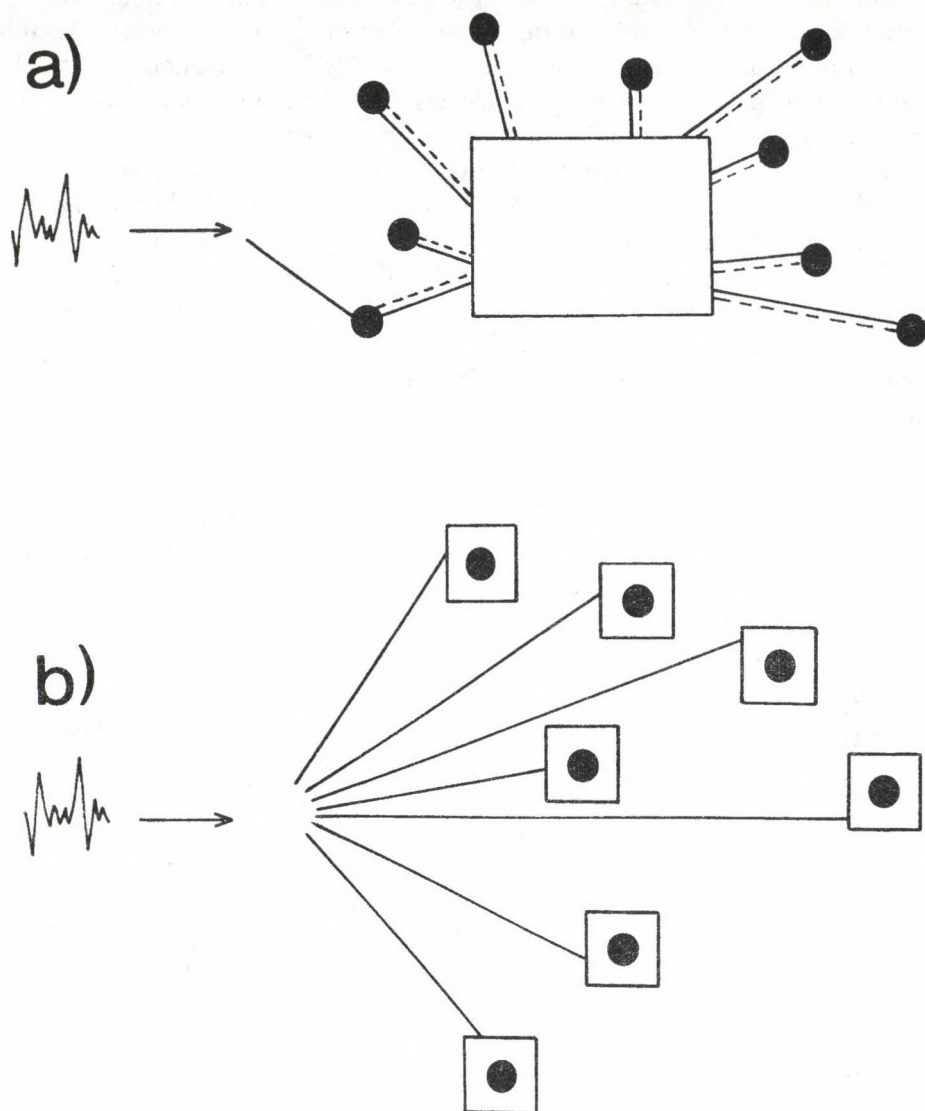


Felmerül a kérdés, hogy vajon milyen a mentális lexikon szerveződése: szemantikai egységeket és azok összefüggéseit tartalmazza (jelentésmezők), avagy a szemantikai egységek és a relációk együttesen vannak kódolva benne. Példával illusztrálva: az egyik feltevés szerint ez azt jelentené, hogy a mentális lexikon egység formájában tartalmazza például a nyelvre jellemző grammatikai szabályokat is (*adom, adod, adja* stb. az *ad* jelentés alatt), míg a másik feltevés szerint a szemantikai egységek és a relációk szoros együttműködésben, de külön vannak reprezentálva (pl. az *ad* fogalomkör, illetőleg a tárgyas ragozás paraméterei). Nem zárható ki, hogy a mentális lexikon szerveződése sokkal jobban nyelvspecifikus, mint gondolnánk. A leírtakból egyértelműen következik, hogy a lexikális hozzáférés vizsgálatának mind az izolált előfordulása (nem feltétlenül izolált ejtésű!), mind a folyamatos szövegben előforduló szavakat tekintetbe kell vennie. A lexikális hozzáférés kutatásakor el kell különítenünk egymástól a kétféle folyamatot, de lényeges jellemzőket adaptálhatunk az egyik megismeréséből a másikra.

A lexikális hozzáférést reprezentáló modellek két nagy irányzatra oszthatók: az ún. moduláris és az interaktív hipotézisekre (vö. Frauenfelder – Komisarjevsky-Tyler 1987). Természetesen számos "közbenső" elképzelés is ismeretes, amelyek e két alaphipotézis különféle keveréséből alakultak ki (ún. hibrid modellek). A két irányzat (vö. 2. ábra) megnevezésére újabban más kifejezéseket találunk: aktivációs (pl. Morton logogen teóriája) és lexikális kereső modellek (pl. Marslen-Wilson cohort elmélete: Marslen-Wilson 1989).

Mínthogy ezek az elméletek is folyamatosan bizonyos módosulásokon (fejlesztéseken) esnek át, hiszen az újabb és újabb kísérleti eredmények mind pontosabb modellálást tesznek lehetővé; a két alapirányzat is változásokon megy keresztül. Ezek az újrafogalmazások nemegyszer megnehezítik az eltérő kifejezésekkel címkézett teóriák azonosítását (vö. például Zwitserlood felosztását: 1989). Ennek ellenére világosan látszik, hogy a moduláris hipotézis újjáformálódása érhető tetten az aktivációs elméletekben és az interaktív hipotézist látjuk viszont a lexikális kereső modellek címszáva alatt. Mi jellemzi e kétféle irányzatot? A **moduláris hipotézis** lényege, hogy a lexikális hozzáférés folyamatában a műveletek nem (vagy csak igen kevésbé) kapcsolatosak egymással, párhuzamos működések jellemzik a folyamatot, s az egyes szintek kezdetén és végén ún. modulok (innen az elnevezés) végzik az elemzéseket, ezek kommunikálnak, de "vakok" egymás működése és működési eredménye(i) iránt (annál is inkább, mint-hogy a műveletsorok lefolyását időben párhuzamosnak tételezik fel). Az **interaktív modellek** lényege, hogy az egyes szinteken történő műveletek állandó kapcsolattartásban zajlanak, ily módon állandó, "real-time" lehetőség van az együttműködésre, illetőleg a korrekcióra (ezek a modellek lehetővé tesznek párhuzamos és sorozatos működést is az interakciók révén).





2. ábra  
A szófelismerési modellek két alaptípusa

Mindkét elmélet számos kísérleti adaton alapszik, s nagyjából azonos mértékűek a bizonytalansági tényezőik is. Nézzük meg, mi jellemzi az aktivációsnak nevezett modelleket! Elsősorban a párhuzamos összehasonlítások, ami azt jelenti, hogy minden egyes szó (a mentális lexikonban) rendelkezik az összehasonlításhoz szükséges jegyekkel és stratégiákkal, és ezek párhuzamosan dolgoznak a bemenő (akusztikai) jel folyamatos feldolgozásának megfelelően. A modulok tehát – amennyiben ezt a kifejezést használjuk – maguk a szemantikai/lexikai egységek, amelyek a műveletek központjai. A lexikális kereső modellekre – csakúgy, mint az interaktív modellekre – a sorozatos összehasonlítások a jellemzőek, amelyek végeredménye egy központi "komparátorban" összegződik, amely mind az inputtal, mind a szemantikai/lexikai egységekkel kapcsolatban van (vö. 2. ábra). Az első feltevés szerint tehát a szavak önmagukat hasonlítják a bejövő jelhez, míg ez utóbbiban a központi komparátor végzi ezt a feladatot: a szavak egy meghatározott csoportjával veti össze a bejövő jelet. Az akusztikailag igen hasonló jelsorozatok (pl. *bor* és *por* vagy *garázs*, *darázs*, *parázs*, *varázs*) mindkét teóriában komoly nehézséget jelentenek. A lexikális kereső modellekben a gyakoriság tényezőjének kiemelésével próbálnak megoldást találni, míg az aktivációs modellekben az aktiváció "súlyának" megváltoztatásával, egyfajta küszöbérték kijelölésével igyekeznek az ilyenfajta megkülönböztetéseket magyarázni. (A probléma megoldatlansága azzal függ össze, hogy mind a "gyakorisági tényező", mind az "aktivációs súly" inkább elméleti, mint objektíven definiálható és igazolható kritériumok, ezért nincsenek kétséget kizáró kísérleti vagy tapasztalati adatok a működésükre vonatkozólag.)

A lexikális hozzáférés folyamatát általában két részre bontják (noha e két rész kísérletileg adatolt elkülönítése lényegében megoldatlan); s számosan mellett foglalnak állást, hogy nincs értelme a két részre bontásnak. A **kétszakaszos hipotézis** szerint az első szakaszban azok a folyamatok mennének végbe, amelyek eredményeképpen az akusztikus hullámformából feldolgozódnak a fonetikai információk, s ezeket a szavakra már korábban tárolt jegyekkel vetjük össze (word recognition). A második szakaszban azokat a folyamatokat tételezik fel, amelyek azért felelősek, hogy a mentális lexikonban a végleges szófelismerés, a "hozzáférés" megtörténjen (lexical access). A két folyamat eredményét azonban igen nehéz (hacsak nem lehetetlen) egymástól elkülöníteni, így kérdésessé válik a folyamat kettébontásának értelme. Elméletileg valóban feltételezhető e két, egymásra épülő folyamat léte. Az első szinten ugyanis létrejön a bejövő jelsorozat és a lehetséges szavak közötti kapcsolat, míg a második szinten bizonyos posztlexikális folyamatok működnek, amelyek révén megtörténik a szójelöltek közül az egyedül lehetségesnek a kiválasztása. Az első folyamat eredménye tehát egy vagy több valószínű szó (pl. *garázs* vagy *darázs*), tehát egyfajta feltevés; míg a második folyamat eredménye az a szó, amelyet kétséget kizáróan felismertünk (*garázs*), tehát egyfajta döntés. Kimondható elméletileg, hogy például az anya-



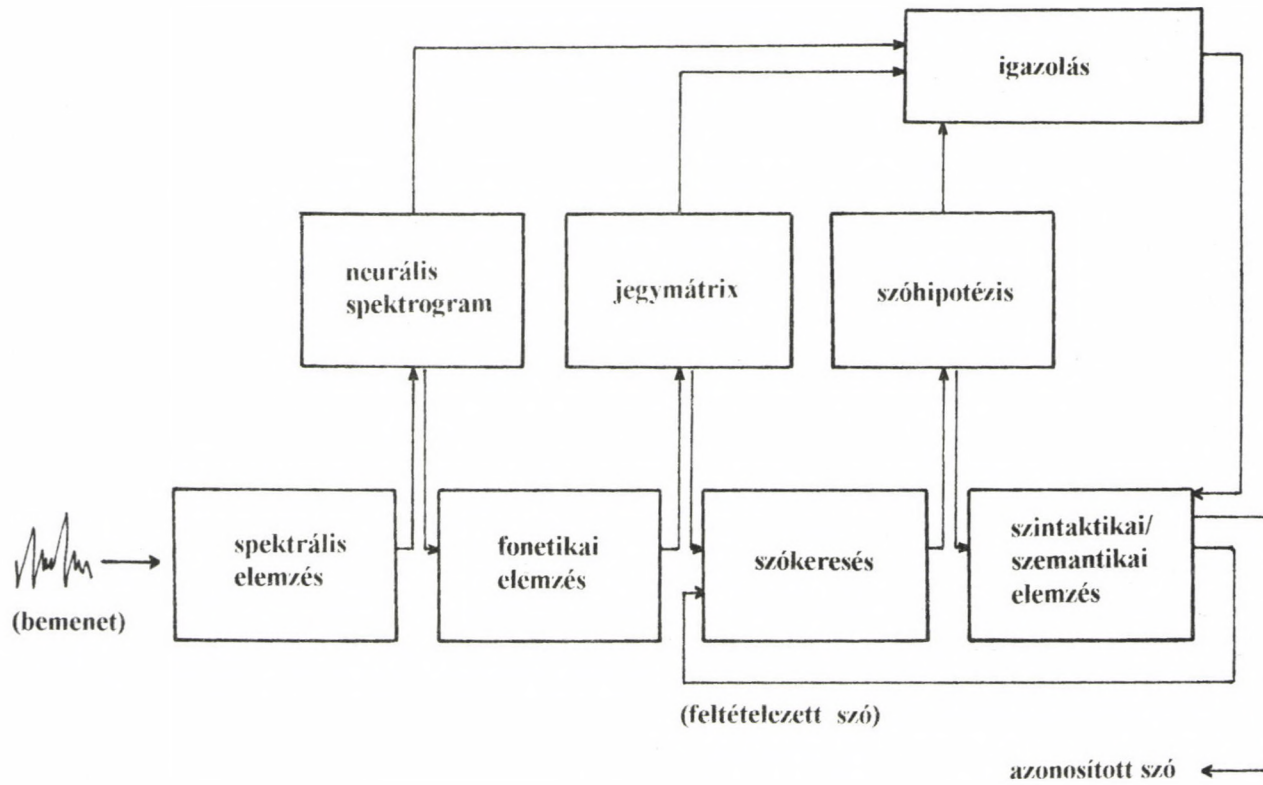
nyelvfejlődés folyamatában az első szakasznak a jelentősége a nagyobb, hiszen a gyermeknek a bizonytalan beszédészlelése és beszédmegértése miatt szinte állandóan készen kell lennie a változtatásra. Annál is inkább, mivel – beszédtapasztalatának csekély volta, illetőleg gyakorlatlansága miatt – sokkal kevésbé tudja a szintaktikai és szemantikai tényezőket felhasználni a szófelismerés folyamatában, mint a felnőtt. Igen gyakran a beszédsszituáció biztosítja számára a pontos szóazonosítást, ehhez pedig nélkülözhetetlen, hogy több szójelölt is "kénél legyen", azaz aktiválva legyen addig, amíg például a beszédsszituációban rejlő invariáns tényező(k) lehetővé teszi(k) a második szakasz létrejöttét, a biztos szófelismerést. A felnőtt beszédmegértési mechanizmusának erre a hangsúlyeltolódásra (az első és a második feltételezett szófelismerési szakasz között) nincs szüksége, mivel többszintű invarianciajelenségek (vö. Gósy 1989) biztosítják a folyamat eredményes működését.

Az elmúlt évtizedekben alkotott sok szófelismerési modell közül kettőt választottunk ki, amelyek működését közelebbről elemezzük: e két hipotézis a fentiekben vázolt elméleti problémák két eltérő megoldási kísérlete.

**1. Az egyik Marslen-Wilson cohort elmélete**, amely a balról jobbra végmenő elemzéseket hangsúlyozza egy szűk fonetikai, illetőleg jegykereső modell keretein belül. A cohort elmélet azt állítja, hogy a szófelismerés folyamatában valamennyi azonosan kezdődő szó aktivizálódik az adott akusztikai input következtében – ez az akusztikai input a "cohort" (< Lat. cohors 'sereg'). Ha a cohort egyetlen szóra redukálódott, akkor megtörtént a felismerés, akkor is, ha a szó teljes egészében nem hangzott el. Ez úgy lehetséges, hogy minden egyes szóra feltételez egy olyan pontot (ez a felismerési pont), amelytől kezdődően a szó más-sal nem cserélhető fel, azaz biztosan megjósolható. Magyar példával illusztrálva: a *macska* szó felismerési pontja a [tʃ] mássalhangzó, mivel nincs olyan szavunk, amely *macs-* kezdetű és *-ka* végű lenne. (A *macska* szó biztos bejósolásához, tehát azonosításához első szótagja tökéletesen elegendő.) A cohort elmélet előnye, hogy jól alkalmazható a folyamatos beszédben történő szófelismerésre, azaz képes tekintetbe venni a kontextushatást is. Hátránya ugyanakkor, hogy nincs egyértelmű magyarázata azon szavak felismerésére, amelyeknél a felismerési pont gyakorlatilag egybeesik a szó abszolút végével, a lecsengéssel, például *le-meZ*, *lemeGY* vagy *majoM*, *majoR*, *MajoS*.

**2. A humán beszédmegértési folyamat valós működésének reprezentálására** Klatt legmegfelelőbbnek az "**analízis szintézissel**" elméletet tartja (1979; 1989). Ennek az elméletnek a lexikális hozzáférést szemléltető modellje pedig a jelen szerző véleménye szerint a legmeggyőzőbben mutatja be a folyamat működését (3. ábra).





3. ábra  
Az "analízis szintézissel" modell

A teória lényege a "készíts hipotézist és ellenőrizd" technika ("hypothesize and test"), amelyet egyébként 1961-ben alkalmaztak először a beszéd formánsainak elemzésével kapcsolatosan (Klatt 1989, 180). Az állítás az, hogy eredményesebb a lexikai hipotézis bejósolása és igazolása az adott szó akusztikai jegyei alapján, mint az akusztikai jegyeket fonetikai vagy artikulációs jegyeknek megfeleltetni, mivel ezek a műveletek rendkívül bizonytalanok. Megjegyezzük, hogy az analízis szintézissel elmélettel szemben két jelentős bírálat merült fel. Az egyik az, hogy nincs egyértelmű magyarázata arra vonatkozólag, hol kezdődik a szöfelismerési folyamat a modellben, azaz hogyan jön létre az elsődleges "lexikális hipotézis". A másik az, hogy túl sok a kognitív szintre utalt művelet, és túl sok az időegységre eső szekvenciális döntések száma a modellben. Az első bírálat ellenérve az, hogy az elsődleges lexikai hipotézis létrejötte lehetségesnek látszik egyfajta durva akusztikai analízis felhasználásával. Ily módon a modell a lexikális hozzáférés folyamatában három szakaszra osztható, ezáltal a kognitív szintre utalt működések száma magától értetődően korlátozódik (második bírálat). A három szakasz a következő: (i) elsődleges és általános fonetikai osztályozás az adott akusztikai hullámforma durva analízisének alapján; (ii) lexikális keresés a lexikonban az első szakasz adatai alapján egy (remélhetőleg) már korlátozott számú szömenyiség kijelölésére; (iii) a kijelölt szavak közti válogatás egy részletes analízis szintézissel igazolási folyamat segítségével, amely a végső szöfelismeréshez vezet.

A folyamat, részletezve, a következőképpen zajlik. A beszéd hullámformája az első állomáson spektrális elemzésen esik át, amelynek eredményét a rendszer a tárolási szinten a neurális spektrogrammal veti össze. Ez az összehasonlítás továbbítódik a legfelső szintű igazolási folyamatokhoz, míg ugyanakkor a folyamat az összehasonlítási eredményeket továbbítja a következő szintre, a fonetikai elemzésekhez. Innen a tárolási szint jegymátrixába vezet az út, ahol ismét összevetések történnek, s folytatódik a kettős továbbítás: egyfelől az igazolási folyamatokhoz, másfelől tovább a szókeresési stratégiákhoz. Így egyidejűleg történik a minden szinten lejátszódott folyamatok igazolása/értékelése, valamint a folyamat csaknem hierarchikusan megoldott továbbműködése, a mind célzottabb elemzés. A szókeresési stratégiák vezetnek a lexikai hipotézishez, majd onnan ismét egyfelől az igazolási folyamatokhoz, másfelől pedig a szintaktikai/szemantikai elemzésekhez. A különböző szintekről jövő működéssorozatok igazolási folyamatairól nincs közvetlen visszajelzés a forrás felé, csak a teljes folyamatműködés végén, a szintaktikai/szemantikai elemzésekhez csatlakoznak be az igazolási műveletek. Ennek nyilvánvaló oka az, hogy az egész mechanizmus egyfajta hipotézisrendszerben dolgozik, s a kettős ellenőrzés összevetése és összegzése elegendő a legvégző ponton. Annál is inkább, mert az elmélet az utolsó működéseknel még egy "biztosítékot" épít a rendszerbe: a szintaktikai/szemantikai elemzéseket követően a folyamat visszakapcsolható a szókeresés (korábbi) szintjére,



s a lexikális hipotézis az összes korábbi adat birtokában újraindítható, vagyis az összes elemzett információ birtokában (számítógépes szakkifejezéssel élve) "bútolható". Az output tehát többszörösen biztosítva van: a hierarchikus jellegű lexikális felismerést az ugyancsak többszörös igazolási folyamatok segítik, és a folyamat ezen végső szakasza újraindítható, vagyis csaknem korlátlan a korrekció lehetősége. Ez az elmélet jól alkalmazható torzított vagy külsőleg zavart akusztikai jelsorozat feldolgozására is (bár az eredetiben erről nincs szó). Minden további nélkül feltételezhető ugyanis, hogy az alsó szintek csak az igazolási folyamatok azonnali (pozitív) visszajelzése esetén lépnek működésbe – természetesen a megfelelő tárolási és összevetési folyamatok közbeiktatásával. Ez ugyan nyilván lassítja a rendszert, de biztosítja az azonnali korrekciót és nagymértékben csökkenti a hibalehetőséget. Felmerül, hogy vajon a legfelső szintű igazolási folyamatok elhagyhatók-e, azaz létezik-e olyan helyzet, amikor működésük nem szükségszerű. Az elmélet eredendően a folyamathoz tartozónak ítéli a mindenkori igazolási folyamatok jelenlétét; ugyanakkor nincs egyértelmű bizonyítéka arra nézve, hogy ezek nélkül a folyamatműködés bizonytalanná, esetleg lehetlenné válna. (Az elmélet jól reprezentálja a különféle szófelismerési zavarok kialakulásának lehetséges okait és helyeit /beszédszlelési nehézség, afázia egyes fajtái stb./, amikor az elemzett folyamatok és működések hiánya, illetőleg bizonytalansága következtében lépnek fel a nehézségek.)

## Kísérleti eredmények

A lexikális hozzáférés folyamatának sajátosságait a magyarban kísérletsorozatokkal vizsgáltuk. Az elsőben arra kerestünk választ, hogy a **kontextushatás miként érvényesül a szókapcsolat szintű szóértésben** az izolált szavak felismeréséhez képest. A második kísérletsorozatban ennek a **kontextushatásnak** egy sajátos változatát vizsgáltuk meg, az **időfaktor sajátos funkcióját**.

### I. A kontextushatás a szófelismerésben

Kísérleti anyagunkat 6 magyar anyanyelvű férfi spontán beszédanyagából választottuk ki. Valamennyien köznyelvi beszélők voltak, hosszabb ideig külföldön nem éltek (budapestiek), nem beszédhibások. Beszédüket tekintve az alapvető különbség a beszédtempójuk volt, amely a 7,4 hang/s-os átlagtól a 16 hang/s-os átlagig változott. Az artikulációs tempójuk átlagértékének határértékei: 8,3 hang/s – 17,1 hang/s. 28 szókapcsolatot, illetőleg ebből 28 szót választottunk ki, amelyeket a 700-as típusú laboratóriumi hangspektrográf segítségével izoláltunk úgy, hogy az elektromos kivágás következtében kattánás vagy egyéb zaj ne előzze meg és ne kövesse a kivágott beszédrészt. A szókapcsolatok kiválasztásának szempontjai a következők voltak: a) különféle szófajú szavak forduljanak elő (*csak, átszaladunk, szorgalom, jobban* stb.); b) egy- és többszótagú szavak



egyaránt legyenek (*megpróbálják, gondolni, voltam, járt* stb.); c) különféle toldalékkal ellátott szavak szerepeljenek (*dolgozott, könyvekben, oka, magasabb, mondjuk, lehetőséget* stb.); d) különböző gyakoriságúak legyenek (*lehet, tehát, mentem, általános, adatokból* stb.); e) szótári és nem szótári előfordulásúak legyenek (*magyar, külön, júniusba, tanítják* stb.). Minden szókapcsolat 2 szóból állt, amelyek közül vagy az első vagy a második tag volt a keresett szó, amelyet az izolált szóteszthez választottunk ki a fentiekkel megegyező kritériumok alapján. A szókapcsolat két tagja között vagy szintaktikai, vagy szemantikai kapcsolat (vagy mindkettő) volt, bár a legtöbb esetben lehetetlennek mutatkozott ezek egyértelmű szétválasztása.

F-J intenzitás- és alaphfrekvencia-mérővel meghatároztuk a keresett szavak és a szókapcsolatok időtartamát. Egy korábbi kísérletünk eredményei alapján (Gósy 1991a), a keresett szavak artikulációs tempójának értékeit három csoportba osztályoztuk: lassú, közepes és gyors minősítéssel (1. táblázat).

Véletlenszerű sorrendben magnetofonra rögzítettük a 28 (mesterségesen) izolált szót, majd a 28 szókapcsolatot (amelyek tehát első vagy második tagként tartalmazták a keresett szót). Négy kísérleti csoportnak játszottuk le 4 hét különbséggel a kétféle hanganyagot: az első csoportban 24 négyéves óvodás, a másodikban 24 ötéves, a harmadikban 24 hatéves óvodás gyermek vett részt; a harmadik csoportot 25 felnőtt (férfi és nő) alkotta, átlagéletkoruk 30 év volt.

A gyermekek feladata a lejátszott hangsor (szó?) felismerése és ismétlése volt (amit a kísérletvezető lejegyzett); a felnőtteknek az azonosított szót le kellett írniuk.

1. táblázat

Artikulációs tempó	A keresett szavak szótagszáma					
(hang/s)	1	2	3	4	5	
5,8 – 10,4	3	3	2	3	-	
11,4 – 16	-	4	5	-	2	
17,3 – 23	-	3	-	3	-	
Összesen	3	10	7	6	2	= 28

A keresett szavak tempója és szótagszáma

A 2. táblázatban összegeztük a szavak felismerésére kapott eredményeket.

A statisztikai vizsgálatok alátámasztották, hogy az elvárásnak megfelelően szignifikáns az eltérés ( $p < 0,01$ ) a négy-, ötéves gyermekek és a felnőttek szófelismerési teljesítménye között (annál is inkább, hiszen a gyermekek számára a kísérleti szavak egy része ismeretlen volt, tehát értelmetlen hangsorként funkcionáltak). Ugyancsak szignifikáns a különbség ( $p < 0,01$ ) mind a gyermekeknél, mind a felnőtteknél az izolált előfordulású szavak és ugyanazon szavaknak szókapcsolatokban való korrekt felismerése között. Nem volt szignifikáns a külön-

ség a hatévesek és a felnőttek teljesítménye között az izolált szavak felismerésében; ami azt jelenti, hogy hatéves korra a szófelismerési folyamatok eredménye megközelíti a felnőttekét. A kontextus minden életkorban – a feltételezésnek megfelelően – nagy segítséget jelent: ugyanazt a szót könnyebben és pontosabban ismerik fel a kísérleti személyek, ha nem önmagában, hanem egy másik szóval alkotott kapcsolatában hangzik el (4. ábra).

2. táblázat

Kísérleti csoportok	A szavak helyes azonosítása (%)	
	izoláltan	szókapcsolatban
négyévesek	45,5	76,75
ötévesek	51,75	85,75
hatévesek	59,8	83,8
felnőttek	66,14	92,85

Spontán beszédből kivágott szavak felismerése

A kontextushatás már négyéves kortól nagymértékben érvényesül. Választ kerestünk arra, hogy vajon a keresett szó artikulációs tempója mennyire befolyásolja a felismerhetőségét, illetőleg módosul-e a kontextushatás a keresett szó artikulációs tempójának függvényében (3. táblázat).

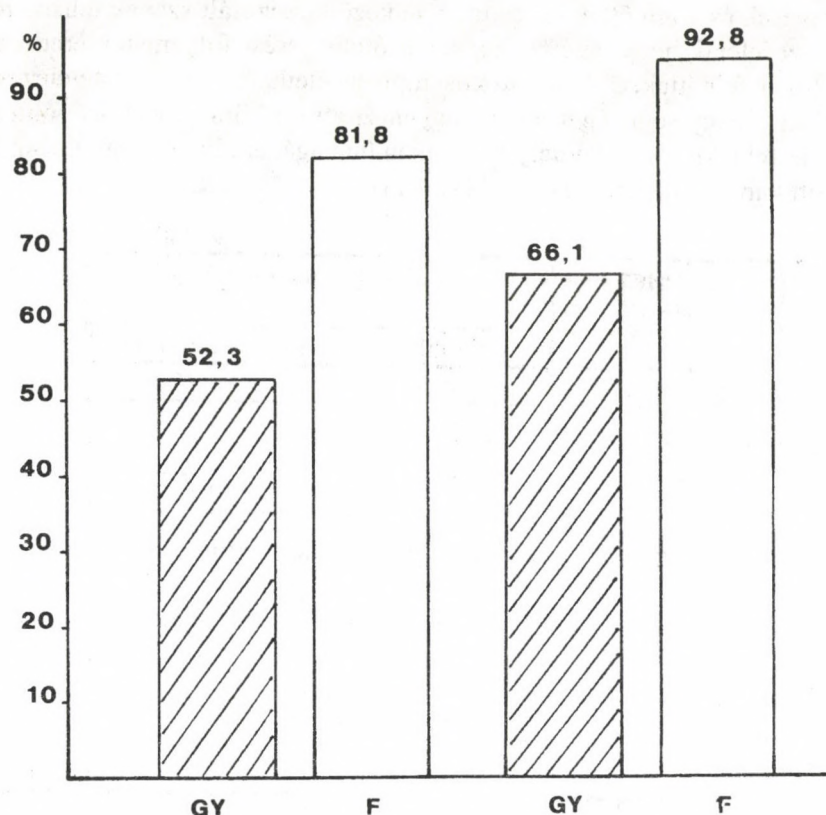
3. táblázat

A szó artikulációs tempója (hang/s)	Helyes szófelismerés (%)			
	izoláltan		szókapcsolatban	
	gyermek	felnőtt	gyermek	felnőtt
5,8 – 10,4	80,3	89	94,7	98,75
11,4 – 16	41,6	62,5	70,4	90,9
17,3 – 23	20,8	41,3	73,6	85,4

Helyes szófelismerés az artikulációs tempó függvényében

Az adatok természetesen igazolták a feltételezést, hogy a helyes szófelismerés nagymértékben függ a szó artikulációs tempójától: minél gyorsabb az elhangzás, annál pontatlanabb az azonosítás, különösen gyermekkorban. Az adatok azonban azt is mutatják, hogy más tényezők ugyancsak befolyásolhatják a felismerést, s ezek működése esetén a tempó másodlagossá válhat. Ilyen például a szó gyakorisága: a *különböző* szó felismerése csak 58,3% a gyermekcsoportban, míg 100%-os a felnőtteknél; vagy a – gyermekeknél, úgy tűnik, gyakoribb – *magyar* szó helyes felismerése náluk 72%, míg ugyanezt a szót a felnőttek átlagosan csak 54%-ban tudták azonosítani.





4. ábra

A kontextushatás érvényesülése gyermekek és felnőttek szófelismerésében

Elemeztük a kapott adatok alapján, hogy a szó jelentésének biztos ismerete hatással van-e a felismerésre. A hipotézisünk szerint a gyermekcsoportok gyakori tévesztése ("gyengébb" teljesítménye) mögött ugyanis a szemantikai információ hiánya húzódik meg. Az adatok ezt a feltételezést csak részben igazolták: lassú tempó esetén alig volt különbség a keresett szó azonosításában függetlenül az ismertségtől és a kontextustól. Gyorsabb elhangzás esetén a szavak azonosítása jobb volt, ha a gyermekek ismerték a szót, de csak akkor, ha a keresett szó a szókapcsolatban fordult elő. Ez azt jelenti, hogy a kontextushatás a szemantikai tényezőkkel együtt érvényesül. Meglepetésként azt láttuk, hogy a gyermekek felismerési folyamata ugyanolyan, sőt biztosabban működik izolált előfordulású szó azonosításakor, ha a szót nem ismerik. Ekkor tehát csak beszédészlelés, a hangsor azonosítása történik meg (és nincs kapcsolat a mentális lexikonnal). Ennek feltételezett oka az, hogy az anyanyelv-elsajátításnak ezen a szintjén a gyermekek észlelési stratégiája hangsúlyozottabb, mint a felnőtteké (vö. készenlét az



új szavak tanulására); s ez teszi lehetővé a pontos hangsor-azonosítást. Hozzájárul ehhez még a gyermekek és a felnőttek közötti stratégia különbsége is. A felnőttek minden esetben igyekeznek értelmes szót felismerni, míg a gyermekek számára a jelentés nélküli hangsorok feldolgozása megszokott (aktív szótanulás korszaka).

Megnéztük, hogy a keresett szó helye a szókapcsolatban meghatározza-e a felismerési biztonságot. Úgy gondoltuk, hogy ha a szókapcsolat második tagja (tehát a később elhangzó) a keresett szó, akkor a felismerés jobb lesz, hiszen az első szó alapján a hallgató már bizonyos predikciókkal rendelkezik. Az adatok ismét csak részben igazolták a feltevést (4. táblázat).

4. táblázat

A szókapcsolat tempója (hang/s)	A szó helye	Helyes felismerés (%)	
		gyermekek	felnőttek
5,8 – 10,4	első	100	100
	második	94,1	98,6
11,4 – 16	első	63,1	91
	második	83,3	90,6
17,3 – 23	első	62,5	81,2
	második	95,8	93,7
Átlagok	első	68,6	89,42
	második	91,2	95,83

#### A keresett szó helyének szerepe a felismerésben

A gyermekek számára sokkal meghatározóbb a keresett szó helye a szókapcsolatban, mint a felnőttek számára. A gyorsabb artikulációs tempójú szavak esetében szignifikánsan jobb a szóazonosítás, ha a keresett szó a szókapcsolat második tagja, vagyis ha a keresett szó egy elhangzó szót követ. A felnőtteknél ilyen különbséget csak az igen gyors szavak felismerésekor láttunk; lassabb tempó esetén a szókapcsolatban elfoglalt hely nem mutatkozott meghatározónak. Az előző adatokkal összevetve, ez azt jelenti, hogy a kontextushatás az adott szó pozíciójától függetlenül érvényesül, tehát a lexikális hozzáférés folyamatában ekkor nem vagy nem elsődlegesen a predikció, hanem más tényező játszik szerepet a felismerésben. Ugyanígy nehéz meghatározni a szemantikai és/vagy szintaktikai hatások megjelenését a szókapcsolat első szavának felismerésében (hiszen azok a keresett szó jelentős részének feldolgozása után jöhetnek csak számításba, s döntően a végső elemzéseikig csak mint jósolt, feltételezett viszonylatok érvényesülnek). Magyaratzként a következő **hipotézist** állítottuk fel. Az analízis szintézissel elmélet kereteiben gondolkodva, feltételeztük, hogy a hallgatónak nem szükséges a keresett szóról azonnali döntést hoznia, ha tudja, hogy van "ide-

je" a végső folyamatok újraindítására és újraellenőrzésére (és természetesen korigálására). A modell lehetővé teszi, hogy a mechanizmus úgy működjön, hogy az első szóról hozott végleges döntést megelőzően bizonyos elemzések megkezdődjenek a második szóra vonatkozólag. Ekkor egyfajta mesterséges várakoztatás történik, a lexikai hipotézis szintje aktiválódik, s mindaddig aktiválódott állapotban marad, amíg a rendszer valamilyen jelzést nem kap arra vonatkozólag, hogy a lexikális hozzáférés folyamatát be kell fejeznie, döntést kell hoznia. Ebben a mesterséges várakoztatási állapotban a rendszer más szintjein történhetnek működéssorozatok; hangsúlyozandó azonban, hogy az első lexikális hozzáférési folyamat ekkor még nem zárult le. Ha pedig ez bizonyíthatóan így van, akkor kimondható, hogy nem a részben párhuzamosan futó (és egy másik szóra vonatkozó) elemzések információinak feldolgozása segíti a hallgatót a keresett szó felismerésében, hanem az a tény, hogy a szóra vonatkozó döntés, a **végső szóazonosítás késleltetve van**. Ebben pedig az időfaktornak jut a meghatározó szerep: ha megfelelő idő áll a mechanizmus rendelkezésére ahhoz, hogy a végső folyamatokat ismételtesse, újraindíthassa, akkor a döntési eredmény várhatóan jobb, pontosabb lesz. Ennek kísérletes igazolására teszt sorozatot terveztünk és folytattunk le magyar anyanyelvű felnőttekkel.

## II. Az időfaktor mint kontextus

A megelőző kísérlet tapasztalatai azt sejtették, hogy az időfaktor, például a hangos szünet egy fajtája kontextusként működhet. E kísérletben ezt a lehetőséget vizsgáltuk meg a magyarra jellemző [ə] hezitációs szünetjelenség felhasználásával. Köznyelvet beszélő, férfi adatközlő spontán elbeszéléséből választottunk ki 15 szót, amelyek a későbbi teszt anyagát szolgáltatták. A szavak kiválasztásának szempontjai megegyeztek az előző tesztanyagéval. A választott szavak a következők (zárójelben a szókapcsolat második tagja): *igék: veszek (elő), jöttek (ismét), sügött (neki); főnevek: utat (nyit), verset (írni), kecske (füle), mezőn (várt), betűk (írása), radírt (kapott); igenév: szaladni (kell); melléknév: másik (baján); számnév: sok (gyalogost); névmás: amit (gondol); határozószók: már (soha), ró-luk (alig)*. Látható, hogy a keresett szó mindig a szókapcsolat első tagja, amely a legkülönbébb szintaktikai/szemantikai kapcsolatban áll a követő szóval. Az eredeti adatközlővel véletlenszerű sorrendben magnetofonszalagra rögzítettük a szavakat úgy, hogy a bemondó igyekezett normál "szókapcsolatot" ejteni, amelynek első tagja a keresett szó, második tagja pedig a magyarra jellemző hezitáló hang, az [ə] volt (imitálva azt a szituációt, amikor a beszélőnek kódolási nehézsége van, például keresi a megfelelő kifejezést). Ily módon olyan mesterséges szókapcsolat-utánpótlásokat hoztunk létre, amelyeknek teljes időtartama megegyezett az eredeti szókapcsolatok átlagával, s a hezitációs jelenség természetessé tette a kifejezést. A bemondott szavak tempója hasonló volt, lényegesen szűkebb



skálán mozgott, mint a megelőző kísérlet spontán szavai. A határértékek a következők: 10,6 hang/s – 13,7 hang/s. (A bemondótól a lehetőségekhez képest a legtermészetesebb ejtést kértük.)

Kétféle tesztanyagot készítettünk: a) az első a keresett szavakat tartalmazta azoknak a hezitációs [ə]-vel alkotott "szókapcsolatában" (pl. *sok..öö*); a másik b) csak a keresett szavakat tartalmazta, a hezitációs hangrészt elektromos úton levágtuk az eredeti kapcsolatból. Mindkét hanganyagot ún. fehér zajjal fedtük el a felismerés megnehezítésére. A zajjal történő elfedés technikája jól ismert a szakirodalomban; 1951-ben Miller, Heise és Lichten számolnak be e technikával végzett kísérletükről. Eredményeik szerint a kísérleti személyek a mondatban elhangzó szavakat – a háttérzaj ellenére – pontosabban azonosították, mint ugyanazon szavakat izolált előfordulásban (vö. saját első kísérletünk). Az általunk használt jel/zaj viszony –8 dB volt átlagosan. Az egyes szavak között mintegy 25 mp (zajos) szünetet hagytunk a szó felismerésére és rögzítésére.

A kísérleti személyek felnőttek voltak, átlagéletkoruk 35 év (magyar anyanyelvű férfiak és nők), összesen 18 fő. Négy hét különbséggel hallották a kétféle hanganyagot úgy, hogy 9 fő a teljesen izolált szavakat hallotta először, míg a másik 9 kísérleti személy a hezitációs hanggal végződő szavakat először. A kísérleti személyek feladata a hallott szavak azonosítása és helyesírásban történő rögzítése volt. A kísérletet csendesített teremben fülhallgatón keresztül végeztük; minden esetben egyénileg teszteltünk.

Az 5. ábra a 15 szóra kapott helyes felismerés százalékos eredményeit tartalmazza. Az egyes oszlopok szemléltetik, hogy a hezitációs hanggal "megtoldott" szavak felismerése minden esetben jobb volt, mint amikor ugyanezeket a szavakat ugyanolyan torzítás mellett teljesen izoláltan kellett felismerni. A két hanganyagra kapott átlagértékek a következők: a 15 szót a kísérleti személyek teljesen izoláltan 43,3%-ban ismerték fel helyesen, míg az [ə] hanggal "meghosszabított" 15 szó felismerésének átlaga 65,5%. A két kísérletből kapott adatok szignifikáns különbséget mutatnak ( $p < 0,01$ ).

A teljesen izolált szavak között két olyan fordult elő, amelyeket a kísérleti személyek képtelenek voltak azonosítani (*mezőn* és *veszek*); a másik tesztanyagban a *veszek* szót 22,2%-ban felismerték, a *mezőn* szó itt sem volt azonosítható. A válaszok azonban még ekkor is jeleznek minőségi különbséget a két anyagban: a teljesen izolált szót többnyire felismerhetetlennek minősítették, míg a "*mezőn...ö*" elhangzás után általában hasonló struktúrájú szavakat írtak, mint *mondom, majom, mellyel, mozdul, megjön, mozdony*. Ennek a kísérletnek az adatai is megerősítik azt a korábban már leírt véleményünket (Gósy 1991b), hogy a szófelismerést alapvetően az akusztikai szerkezet határozza meg: ugyanazon szavak felismerése ugyanolyan arányban történik a két hanganyagon belül. A részt vevő beszédhangok, illetőleg hangkapcsolódások akusztikája következtében jóval biztosabb az azonosítása például a *már* (100%, ill. 94,4%), a *róluk* (77,7%,

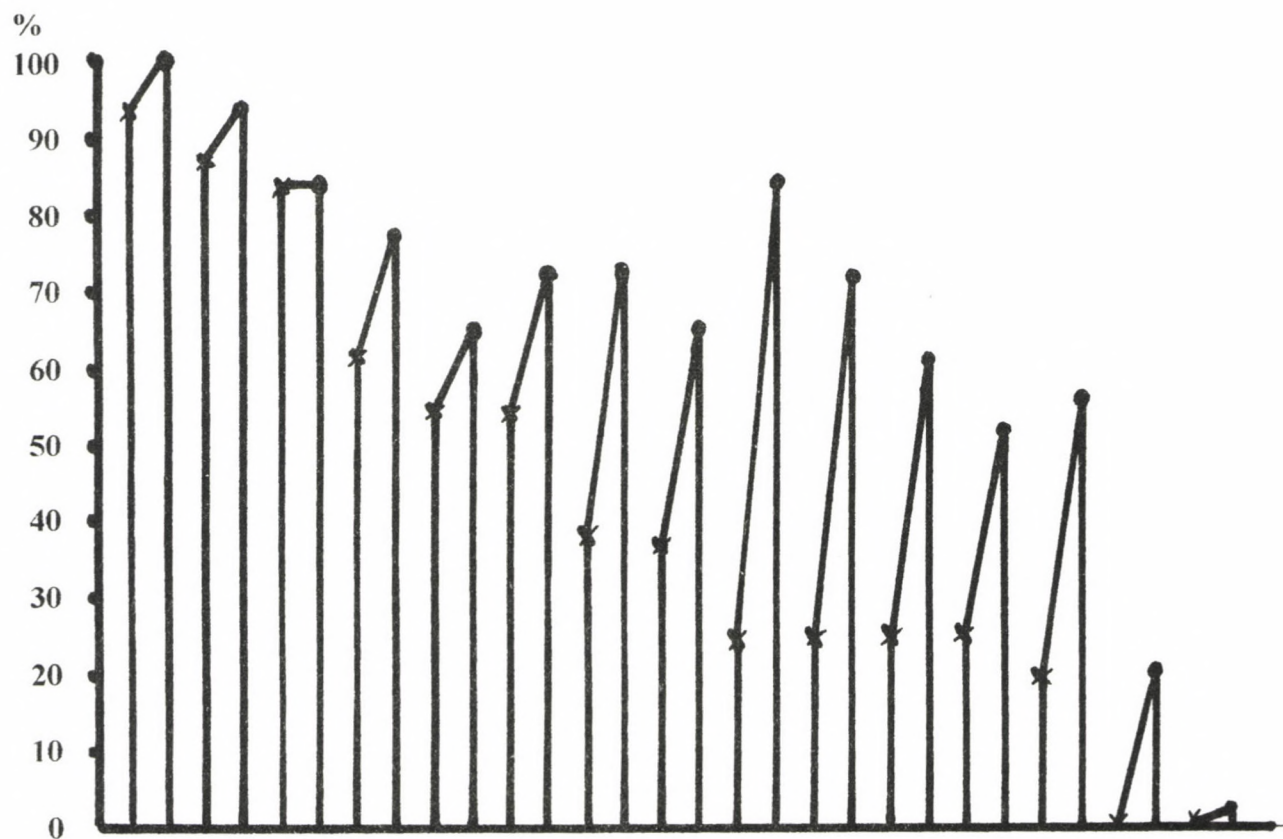
ill. 61,1%), a *másik* (83,3%, ill. 73,3%) vagy a *súgott* (66,6%, ill. 55,5%) szavaknak, míg lényegesen nehezebb például a *veszek* (22,2%, ill. 0), az *amit* (55,5%, ill. 22,2%) vagy a *radírt* (50%, ill. 27,7%) szavaknak. Nincs értékelhető különbség a felismerés eredményeiben a szavak gyakoriságát tekintve (ez részben összefügg magával a felismerési feladattal is, vö. Brown 1990, 23–26).

### Következtetések

A kísérlet eredményei egyértelműen igazolták azt a hipotézist, hogy a szünet – jelen esetben a hangos szünetként funkcionáló hezitálás – kontextusként működik. Ezáltal a lexikális hozzáférés analízis szintézissel modellje újabb megerősítést nyert, hiszen e modell működésével a szünetkontextus egyértelműen megmagyarázható. Talán még ennél is lényegesebb az a tény, hogy a kontextus fogalomkörét kiszélesíthetjük egy további tényezővel, az időfaktorral, amit az első kísérletsorozat egyes adatai valószínűsítettek, a második kísérletsorozat eredményei pedig igazoltak. Kontextushatáson tehát nemcsak a szintaktikai/szemantikai kapcsolatokat, hanem az időtényezőt is értjük. Ez azt jelenti, hogy ha a hallgatónak elegendő ideje van például egyetlen szónak a felismerésére abban az értelemben, hogy a lexikális hozzáférés folyamata úgy kezd működni, mintha a szónál nagyobb nyelvi egység feldolgozására állna be, akkor az izolált szó felismerési folyamata olyan lesz, mint amikor a valódi kontextushatást biztosító szintaktikai és/vagy szemantikai faktorok hatásával számolunk. Még tovább gondolva, ez odavezet, hogy a kontextushatás mint olyan nemcsak az adott szerkezeti/jelentéstani funkciók bármilyen fokú felhasználását jelenti, hanem az adott lexikális hozzáférés folyamatának módosulását is.

Felvetődik a kérdés, hogy vajon az időfaktor mint kontextus nyelvspecifikus jelenség-e vagy nem. Bár nincsenek kísérleti adatok arra vonatkozólag, hogy javul-e a szófelismerés, ha nem kérünk azonnali választ a kísérleti személyektől (tehát konkrétan van idejük a felismerésre), a feltevésünk az, hogy nem javulna. Elméletileg ugyanis csak valamiféle nyelvi inger hatására jöhet létre a lexikális hozzáférés folyamatának módosulása, amit jelen kísérletünkben a hezitációs jelenség felhasználásával értünk el, s önkényesen neveztünk időfaktornak. A hezitálás jelensége nem, de akusztikai/artikulációs megjelenési formája nyelvspecifikus. Hipotetikusan tehát kimondható, hogy e sajátos kontextushatásnak nyelvfüggetlenül be kell következnie minden lexikális hozzáférési folyamatban.





5. ábra

Izolált előfordulású (x—x) és hezitációs hanggal "megtoldott" (•—•) szavak helyes felismerése

### Irodalom

- Brown, C. M.: Spoken-Word Processing in Context. Nijmegen 1990.
- Forster, K. I.: Basic issues in lexical processing. In: Lexical Representation and Process. Ed.: Marslen-Wilson, W. Cambridge, Massachusetts, London, England 1989, 75–108.
- Frauenfelder, U. H. – Komisarjevsky Tyler, L.: Spoken Word Recognition. Cambridge, Massachusetts, London, England 1987.
- Gósy, M.: The perception of tempo. In: Temporal Factors in Speech. Ed.: Gósy, M. Budapest 1991a, 63–107.
- Gósy Mária: Szavak és mondatok megértése. Magyar Nyelv LXXX VII, 1991b, 151–162.
- Klatt, D.: Speech perception: A model of acoustic-phonetic analysis and lexical access. In: Perception and Production of Fluent Speech. Ed.: Cole, R.A. Hillsdale, New Jersey 1979, 243–289.
- Marslen-Wilson, W.: Access and integration: Projecting sound onto meaning. In: Lexical Representation and Process. Cambridge, Massachusetts, London, England 1989, 3–25.
- Miller, G. A. – Heise, G. A. – Lichten, W.: The intelligibility of speech as a function of the context of the test materials. Journal of Experimental Psychology 41. 1951, 329–335.
- Nooteboom, S. G. – Truin, P. G.: Word recognition from fragments of spoken word by native and non-native listeners. IPO Annual Progress Report 15. 1980, 42–47.
- Tanenhaus, M. K. – Lucas, M. M.: Context effects in lexical processing. In: Spoken Word Recognition. Eds.: Frauenfelder, U. H. – Komisarjevsky Tyler, L. Cambridge, Massachusetts, London, England 1987, 213–235.
- Zwitserslood, C. M. E.: Words and Sentences: The Effects of Sentential-Semantic Context on Spoken-Word Processing. Nijmegen 1989.



# MAGYAR AKCENTUS A FRANCIÁBAN: HÁROM FONETIKAI KÍSÉRLET

**Fagyal Zsuzsanna**

MTA Nyelvtudományi Intézete, Budapest és  
Université de Paris III Sorbonne Nouvelle

## **Bevezető**

Az a magyar anyanyelvű beszélő, aki hosszabb-rövidebb ideig francia környezetben él, tapasztalatból tudja, hogy a franciák igen érzékenyek mindenfajta, a sztenderd (párizsi) ejtismódtól eltérő, ún. "idegen akcentusra". Ezt a megfigyelést a jelenség pszicho-szociolingvisztikai vetületeivel foglalkozva Kassai is megerősíti: "amíg pl. az olaszok szívesen fogadnak minden, anyanyelvükön elhangzó megnyilatkozást, a franciáknál egyenesen a szótár szentesíti a 'parler français comme une vache espagnole' (= kerékbe töri a francia nyelvet; szó szerint: úgy beszél franciául, mint egy spanyol tehén) kifejezést" (Kassai, megjelölés alatt).

Amint azt az idézett tanulmány is megállapítja, az idegen akcentusra való érzékenység sok tényező függvénye. Szerepet játszik az adott társadalom struktúrája, geopolitikai hagyományai, az irodalmi nyelv presztízse éppúgy, mint az anyanyelvi és nem anyanyelvi beszédpartnerek közötti társas viszony jellege. Megállapítható, hogy a francia nyelv esetében az adott tényezők, főként a jelenkor etnikai migrációinak hatására, kifejezetten erős érzékenységet eredményeznek az "idegen" ejtismódokkal szemben.

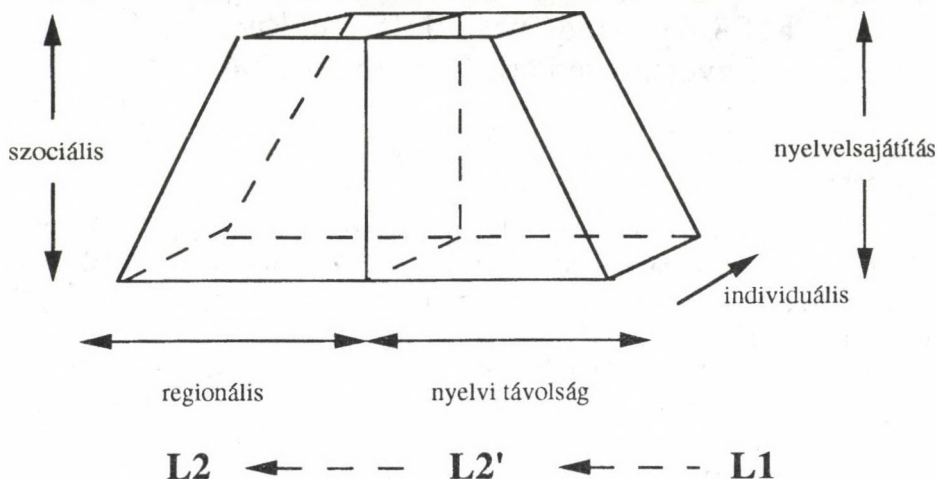
A jelen tanulmány ezt az "érzékenységet" kívánta magyar hanganyagon, néhány kiválasztott szó alapján tesztelni. Olyan mérhető és észlelhető beszédakusztikai jegyek kimutatását tűzte ki céljául, amelyek alapján az anyanyelvi beszélő "akcentusosnak" – tehát nagy valószínűséggel nem született francia által ejtettnek ítéli a hallott szót vagy mondatot.

## **A vizsgálati módszer elméleti modellje**

Az akcentusnak mint nyelvi jelenségnek a vizsgálata számos olyan elméleti kérdést vet fel, amelyet az ezzel foglalkozó, elsősorban beszédkutatások, csak ritkán vizsgálnak behatóbban.

A legnagyobb nehézséget e többváltozós nyelvi rendszer modellezése jelenti, pontosabban az, hogy hogyan képes az adott elméleti vizsgálati modell megoldani individuális stílusjegyek egy nem anyanyelvi nyelvközösségre való vonatkoz-

tatását. Jelen tanulmány svéd beszédkutatók mintájára az 1. ábrán látható – igen alaposnak tűnő – modellt választotta alapjául (Bannert et al. 1982).



1. ábra

Amint az az ábrán is látható, a svéd modell az adott nyelvi jelenséget egyrészt az anyanyelvi beszélő által képviselt célnyelv (L2), másrészt az idegenajkúak által beszélt célnyelv módosult változata (L2') szempontjából vizsgálja. Minden esetben négy változót különböztet meg, amelyek közül három, a nyelvvelsajátítás mértéke, a szociális, illetve individuális tényező, mindkét nyelvhasználatot jellemzi. A regionalitás a célnyelv sztenderd változatánál (L2), míg az ún. "nyelvi távolság" az idegenajkú által beszélt L2'-nél lép be a modellbe.

Ez utóbbi leginkább azért figyelemre méltó, mert – noha kizárólag fonológiai összehasonlítás alapján<sup>1</sup> – megpróbálja felmérni az idegen anyanyelvű beszélő anyanyelve (L1) és az elsajátítani kívánt célnyelv (L2) közti nyelvi különbségét.

Feltételezése szerint az idegen anyanyelvű beszélő által a célnyelven elkövetett "hibák" anyanyelvi interferenciákkal magyarázhatók. Ezért ismerünk fel bizonyos "típushibákat", amelyeket egy adott anyanyelvre tudunk vonatkoztatni.

<sup>1</sup> "It takes only to consideration two factors: those deviant features in pronunciation that have been judged as having severe consequences for intelligibility and the distribution of these deviant features over the three areas of phonology, namely vowels, consonants and prosody" (Bannert et al. 1982, 33.)



## Vizsgálati feltételek

A bemutatott elméleti modell alapján határoztuk meg a jelen tanulmány kísérleti feltételeit is. Éppen a változók nagy számára való tekintettel e feltételek igen kötöttek voltak.

A kísérletben részt vevő személyek kiválasztásakor az individuális, szociális és elsajátítási változók (ld. 1. ábra) lehető legjobb rögzítése volt a cél. Harminc (15 magyar és 15 francia) 22 és 29 év közötti egyetemistát a nyelvelsajátítás foka, ill. a francia nyelvhez való kapcsolatuk alapján két csoportra osztottunk: franciatanárok és nem franciatanárok. A magyar kísérleti személyek mindkét csoportjánál ügyeltünk arra, hogy mindegyikük 1 évnél nem régebben Párizsban tartózkodó, magyar anyanyelvű diák legyen. (A véletlen és a kötött kísérleti feltételek úgy hozták, hogy a franciatanárok csoportjába csak nők, a nem franciatanárok csoportjába pedig csak férfiak kerültek. Ez a különbség azonban – nézetünk szerint – nem eredményez túl nagy egyéni eltéréseket a vizsgált korpuszon belül.)

A célnyelvre (L2) jellemző regionális különbségek kiküszöbölésére a párizsi ejtismódot vettük alapul, így a 15 francia anyanyelvű beszélőt Párizsban vagy Párizs környékén született franciák közül választottuk ki. A magyar és a francia nyelv közti "nyelvi távolság" becslésére e tanulmány nem tesz kísérletet, ellenben a teszt-szavak és teszt-mondatok kiválasztásakor (ld. alább) támaszkodik más szerzők (Kassai 1974; Herman 1984) által is meghatározott, a magyar anyanyelvűek számára nehézséget jelentő pontokra.

## A vizsgálati anyag

A vizsgálati anyag hitelessége és spontaneitása érdekében nem izolált szavakat, ill. mondatokat választottunk, hanem egy rövid, összefüggő szöveget: Alphonse Daudet "Les chèvres de Monsieur Seguin" [Seguin úr kecskéi] című novellájának bevezető bekezdését (ld. Függelék 1.). E szöveget magyarra is lefordítottuk (ld. Függelék 2.). A prozódiai összehasonlíthatóság érdekében a fordításánál az eredeti szöveg szintaktikai szerkezetéhez való legteljesebb alkalmazkodás volt a cél (ld. harmadik kísérlet).

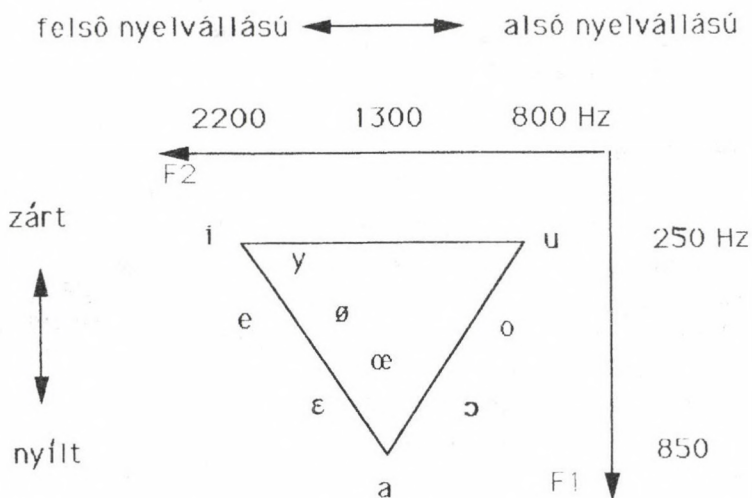
A felvételek a párizsi Fonetikai Intézet csendesített szobájában, Marantz CT130 típusú magnetofon és Vivanco típusú, ruhára erősíthető mikrofon felhasználásával készültek. A magyar beszélők a szöveget először magyar fordításban, majd francia nyelven olvasták el. Mindkét csoport tagjai az adott szöveget először magukban tanulmányozták, majd megszokott beszédtempójukban, természetes hangon és érthetően magnetofonra olvasták. A fenti kiemelés egyrészt a tempógyorsítás elkerülésére, másrészt pedig egyfajta spontaneitás megőrzésére biztatta a kísérleti személyeket. A franciákat azért, hogy a meseként ismert novellát ne "dekklamálva" olvassák, a

magyar beszélőket pedig azért, hogy ne igyekezzenek mindenáron megközelíteni a "jó" francia kiejtést. Ennek ellenére egy spontán utánzás, mint később látni fogjuk, igen sokat elárult a magyar beszélőknek a "tipikus" francia kiejtésről alkotott elképzeléséről (ld. harmadik kísérlet).

### Első kísérlet: a nyílt [ɔ] és a nyílt [œ] magyar kiejtésben

Az első kísérlet egyike azon három kísérletnek, amelyek – célzottan – magyar anyanyelvűek számára kiejtési nehézséget jelentő fonetikai jelenségeket kívánnak tesztelni. A nyílt [ɔ] és [œ] magánhangzók helyes kiejtése eme "neuralgikus" pontok közé tartozik.

A magyar és francia magánhangzó-rendszer összehasonlításakor (Kassai 1971, 1974; Herman 1984) ugyanis kiderül, hogy a magyar nyelvben fonológiai jelentőséggel bíró, de a franciában nem jellemző hosszú-rövid oppozíció, ill. a franciában elsőrendű, de a magyarban csak kísérvjelenséggé mutatkozó nazális artikuláción kívül (ld. 2. kísérlet) a két rendszer közti tipikus különbség az ún. "fél-nyílt" magánhangzók (*voyelles semi-ouvertes*) jelenléte a franciában. A két magánhangzó helyét a francia magánhangzó-rendszerben a 2. ábra szemlélteti.



2. ábra

Itt kell megjegyezni azt a közismert tényt, hogy e két hang a mai modern franciában már nem rendelkezik széleskörű jelentésmegkülönböztető szereppel. Néhány, sokat idézett oppozíció kívül, mint pl. *jeune-jeûne*, [ʒœn - ʒœn] a nyílt [œ] például szinte kizárólag csak zárt szótagban, zárt párjával kiegészítő elosz-



lásban fordul elő. A nyílt [ɔ] fonológiai jelentősége hasonlóképpen "megkopott", bár itt még több minimális párt találni: *Paul-Paule, sol-saule, molle-môle...* stb.

Ennek ellenére, egybehangzó állítások szerint, a két hang helyes kiejtése a "bon usage" (= szép beszéd) egyik legfontosabb ismérve a franciában.

A jelen elemzésben az összefüggő szövegből elkülönített *bonheur* [bɔ̃œʁ] szó szolgált vizsgálati alapul arra nézve, hogy vajon milyen arányban jelent nehézséget a magyar beszélők számára az adott hangkapcsolatokban szereplő két nyílt magánhangzó kiejtése?

A 15 magyar és 7 francia olvasatból kiválasztott *bonheur* szót először a LIMSI-Unice (PC) számítógépes program segítségével kivágtuk a szövegből. Tekintettel arra, hogy a szóvégi uvuláris [R] hang a magyar olvasatokban néha denti-alveoláris "magyar" [r]-ként hangzik el, a szóvégi mássalhangzót is levágtuk. Így olyan hangsort kaptunk, amely ebben a formában mindkét nyelvben megvalósulhat<sup>2</sup>.

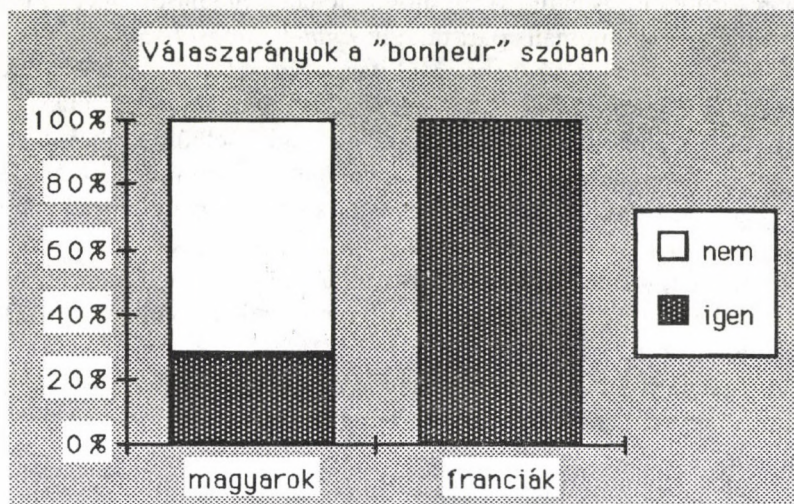
A kivágott 22 szót tetszőleges sorrendben, szavanként 2 ismétléssel, az ismétlések között 1 másodperc, az egyes ejtések között 3 másodperc szünettel bemutattunk 25 francia anyanyelvű, nem nyelvész képzettségű személynek. A tesztszavak meghallgatása zárt helyiségben, fejhallgatón keresztül, egyenként történt. Válaszlapon "igen" vagy "nem" választ kértünk arra a kérdésre, hogy a hallott szavakat francia beszélő ejtette-e ki vagy sem.

### Az első vizsgálat percepciók tesztjének eredménye

A percepciók teszt nem mutatott számottevő különbséget a magyar anyanyelvű franciatanárok és a magyar anyanyelvű laikus beszélők között, bár általánosságban véve megállapítható volt, hogy amennyiben a magyar beszélőt franciának ítélték, úgy az illető beszélő 4 esetből 3-ban franciatanár volt.

Amint azt a 3. ábra mutatja, a magyar beszélőket 73%-ban, 87%-os biztonsággal ismerték fel a franciák. 27%-ukat tartották franciának, köztük 3 tanárt és 1 laikus beszélőt, s ezt 82%-os biztonsággal. A francia beszélőket illetően 100%-os volt a felismerés, de viszonylag nagy szórással ( $\sigma=8,8$ ;  $\text{range}=28\%$ ). A magyarok esetében a francia hallgatók többségének véleményét képviselő nemleges válaszok nagyobb szórást mutatnak ( $\sigma=14$ ;  $\text{range}=36\%$ ), mint a franciák produkcióira adott igenlő válaszok.

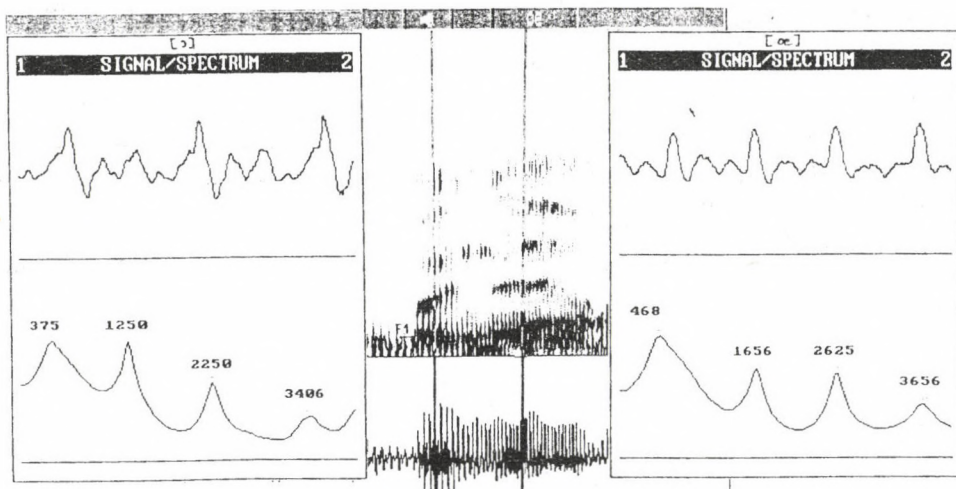
<sup>2</sup> Itt kell megjegyezni, hogy a koartikuláció miatt a nyílt [œ] minősége nem teljesen független az öt követő mássalhangzó uvuláris ill. denti-alveoláris jellegétől, ami - noha nézetünk szerint csak kis mértékben - befolyásolhatja az észleletet.



3. ábra

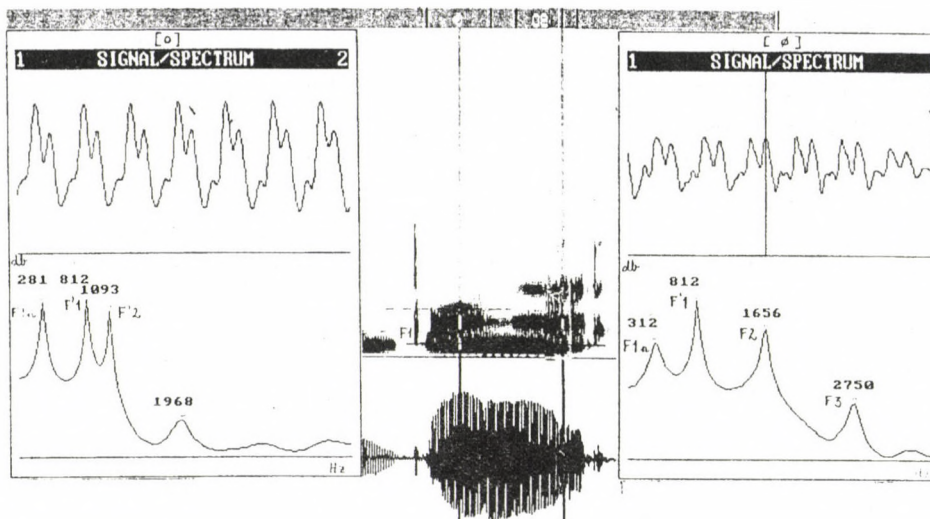
### Az akcentus akusztikai összetevői a "bonheur" szóban

E felismerési arány magyarázatot nyer, amint összehasonlítjuk egy – az esetek 92%-ában 100%-osan – franciának ítélt szó spektrumát (4. ábra) egy másik – az esetek 67%-ban 100%-osan – nem franciának tekintett szóéval (5. ábra). Az első esetben francia, a másodikban magyar női adatközlőről van szó.



4. ábra





5. ábra

A két formánsszerkezetet összehasonlítva jól látható, hogy a francia beszélő esetében mindkét magánhangzó első formánsa (F1) jóval magasabb, mint a magyar beszélő esetében. A szájüreg nagyobb nyitottsága és az ezzel járó, a nyelv mögött képződő üreg térfogatcsökkenése tehát megnöveli az F1 magasságát, ami a francia beszélő esetében megfelel a szakirodalomban megadott formánsértékeknek (Calliope 1989; Carton 1991). Ellenben a magyar beszélő által ejtett, 100%-ban nem "franciásnak" ítélt szó mindkét magánhangzója igen eltérő formánsszerkezetet mutat. A nyílt [ɔ] esetében az első formáns sávszélességének megnövekedése egyértelműen nazalizációt jelez, míg az F2 és F3 formánsértékei csak kevéssel térnek el az orális [o] hangzóra megadott magyar, ill. francia sztenderd értékektől (Calliope i.h.; Carton i.h.; Olaszky 1989).

Az [œ] még ennél is egyértelműbb kettős szerkezetet tükröz: nazális hatást az F1 sávszélességének megnövekedésében, s a két alsó formánsérték tekintetében ( $F_{n1}^3=312$  Hz;  $F1=812$  Hz), és orális magánhangzó szerkezetet a felső formánsok régiójában ( $F2=1656$  Hz és  $F3=2750$  Hz). Az ilyen típusú kettős formánsszerkezetet "nazalizáltak" nevezhetnénk.

Kitűnik tehát, hogy az adott szóban vizsgált nyílt [ɔ] és nyílt [œ] magánhangzók a magyar és francia beszélők esetében sarkalatos eltéréseket mutatnak. A félig nyílt artikuláció a magyar ejtésben nem valósul meg, sőt, egészen más, a francia nyelvben fonológiai jelentőséggel bíró nazális jeggyel helyettesítődik. Az

<sup>3</sup>  $F_{n1}$  = első nazális formáns

okok a magyar kiejtési szabályokban keresendők. Már Bakó is jelzi, hogy "egybeolvad a szótagzáró *m*, *n*, *ny* az előtte álló magánhangzókkal, s vagy teljes orrhangzó keletkezik, vagy pedig csak félig egyesül a két mód" (Bakó 1937, 2). A magyar gépi beszéd szintézis később pedig szabályként rögzíti: "A magánhangzó nazalizálásait úgy oldhatjuk meg, hogy az "u" előtti hangszelet(ek)ben az F1 formáns sáv szélességét fokozatosan növeljük" (Olaszy i.h. 103).

A 100%-osan nem francia ejtésű *bonheur* szóban tehát anyanyelvi interferencia jelenséget tapasztalunk: az intervokális [n] erősen nazalizálja a szomszédos magánhangzókat, akárcsak a magyar beszédben. Ezek formánsszerkezete azonban csak részben nazalizált, a felső formánsok régiói megőrzik orális jellegüket. Így tehát érthető, hogy miért válik tipushibává az orrhangúsodás az *américaine*, *domaine* vagy *semaine* szavakban: úgy tűnik, könnyebb a magyar beszélőnek "nazalizálni" az [n] környezetében, mint annak hatásától függetlenül magát (Földi 1988).

### Második kísérlet: a nazális [ẽ]

Amennyiben az előző kijelentés igaz, úgy nem okozhat(na) problémát a második kísérletben tesztelt *seguin* szó kiejtése, ahol éppen egy nazális magánhangzó, az [ẽ] vizsgálata volt a cél.

A kísérleti eljárás ez esetben is ugyanaz volt, mint az első esetben, annyi eltéréssel, hogy itt a szegmentálási nehézségek<sup>4</sup> miatt csak 12 szót sikerült tesztelni. Mint azt a 6. ábra is mutatja, a teszt eredménye teljesen megegyezik az első vizsgálat eredményeivel, noha a hallott szavak száma, sorrendje, sőt a beszélők személye is nagyrészt különbözött.

A helyes felismerési arány a francia beszélők esetében is 100% volt. A 25 anyanyelvi hallgató 75%-a nemleges, 25%-a igenlő választ adott a magyar beszélők által ejtett szavak hallatán. Ez esetben azonban a francia hallgatók "szigorúbbak" voltak nemleges válaszaikat tekintve, mint a franciák produkcióira adott igenlő válaszaikban. (Az első esetben  $\sigma=14,5$ , range=30%; a másodikban:  $\sigma=16$ , range=48%)

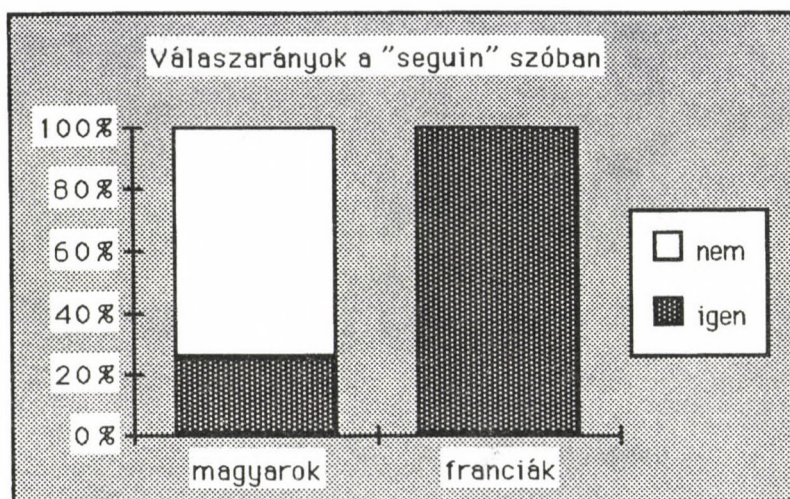
### Az akcentus akusztikai összetevői a "seguin" szóban

A nazális [ẽ] esetében többféle ejtismód figyelhető meg. A 7., 8. és 9. ábrán a három leggyakoribb magyar ejtést tüntettük fel. A 7. ábra az egyetlen, 100%-ban franciának ítélt ejtést szemlélteti. A 8. ábrán látható, 64%-ban inkább külföldinek tartott ejtés, valamint a 9. ábrán bemutatott, 100%-ban nem franciának vélt ejtés 6-6 magyar beszélőnél volt kimutatható. E három spektrum összevethető a

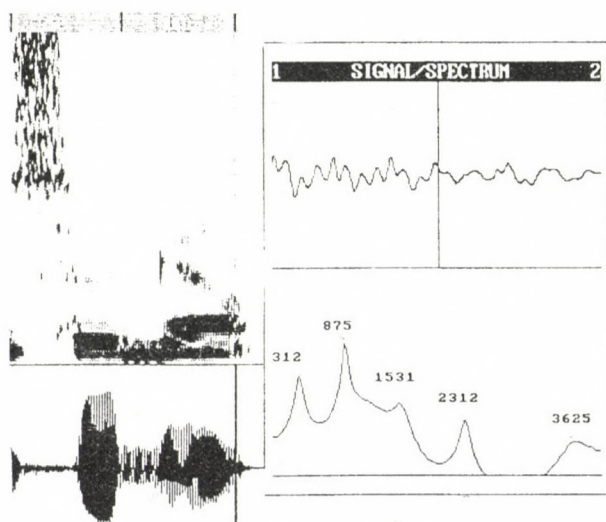
<sup>4</sup> Sok esetben a gyors beszédtempó következtében a nazális magánhangzó időtartama túl rövidnek bizonyult ahhoz, hogy a glottális zár felpattanása után a lágy íny leeresztésével megvalósuló nazális artikuláció tisztán észlelhető legyen.



10. ábrán látható "sztenderd" ejtéssel, amelyet a megkérdezettek 100%-a – helyesen – francia által ejtettnek ítélt.



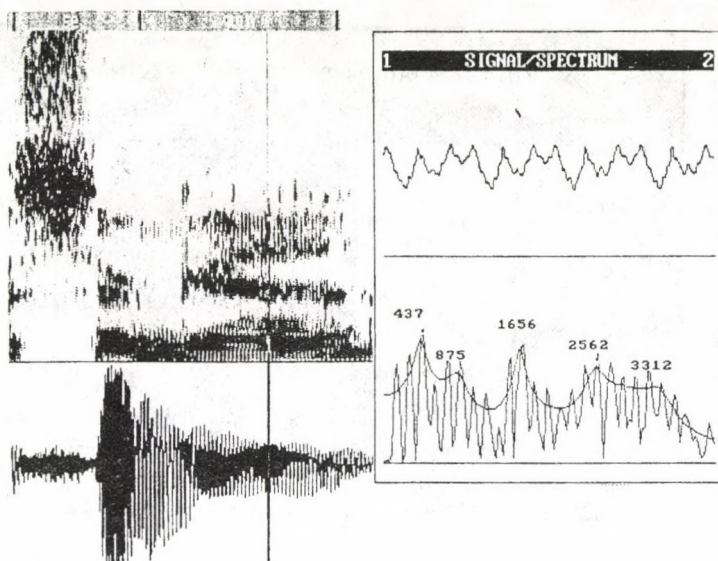
6. ábra



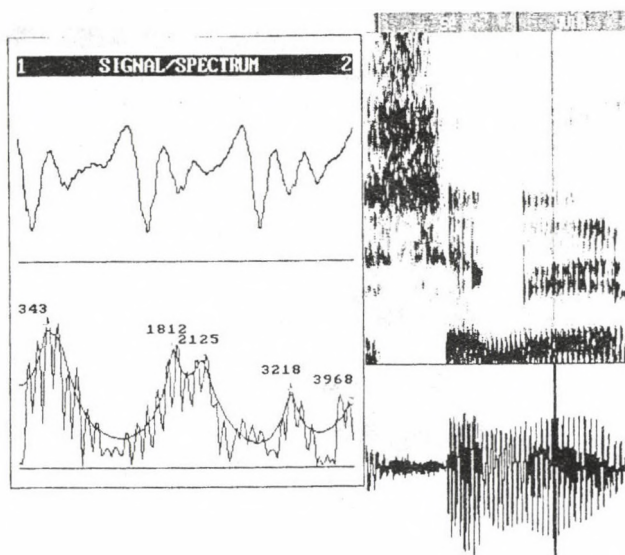
7. ábra

A 7. és 10. ábra összevetéséből kitűnik, hogy a férfi és női hang különbségeiből adódó – némiképp eltérő – formánsértékeket leszámítva a 100%-osan "jó"

magyar ejtés és a 100%-osan "jó" francia ejtés spektrumképe tendenciáit tekintve megegyezik.

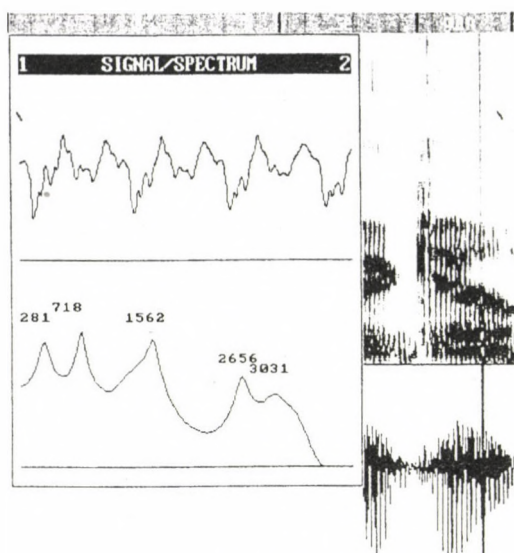


8. ábra



9. ábra





10. ábra

Ezzel szemben az észleletben jelentős változások tapasztalhatók abban az esetben, amikor az alsóbb formánsok régiójában a fent említett szerkezetektől eltérő alakzat valósul meg. A 8. ábrán szemléltetett esetben a válaszok összarányának 64%-ában "nem franciásnak" tűnik az az ejtés mód, ahol az első nazális formáns ( $F_{n1}$ ) és a módosult első orális formáns ( $F1$ ) sávszélessége felcserélődik. Jelen esetben a 437 Hz értékű  $F_{n1}$  túl "magasnak" és túl "intenzívnek" tűnik az  $F1$ -hez képest, amelynek értéke ugyan megfelel a sztenderd méréseknek ( $\approx 800$  Hz), de a sávszélessége túl nagy az  $F_{n1}$ -hez viszonyítva.

Az észlelet szempontjából elsődlegesen fontosnak tűnő alsó formánsrégió "szenvet kárt" akkor is, amikor a mért értékek, illetve a spektrum tisztán orális hangzóra emlékeztetnek (Castelli – Perrier – Badin, 1989; Maeda, 1982).

### A hanglejtés vizsgálata: harmadik kísérlet

A harmadik kísérlet célja az olvasott szöveg néhány mondatának kizárólag a hanglejtés szempontjából történő vizsgálata volt. Jelen esetben a többi változó (pl. időtartam, intenzitás) behatóbb vizsgálatára nem nyílik mód.

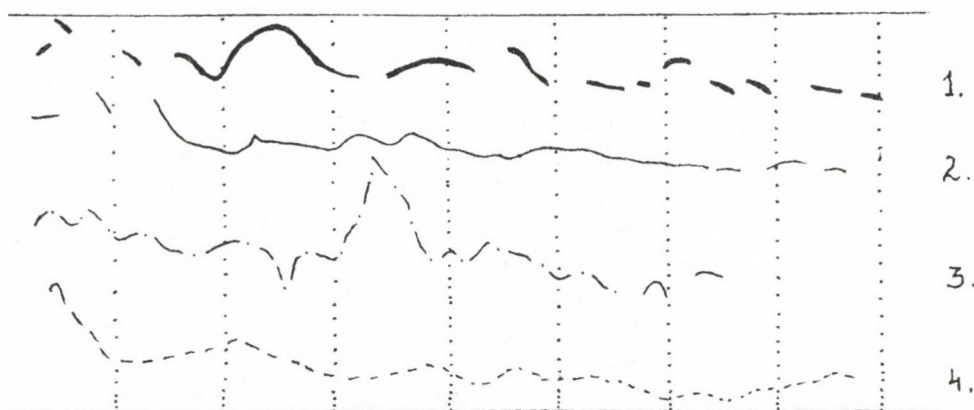
### Megfigyelés

E kísérlet keretében először, megfigyelő jelleggel, összehasonlítottuk a 30 kísérleti személy francia olvasatának első mondatát. A cél annak megfigyelése

volt, hogy hogyan helyezik el a magyar és francia anyanyelvű beszélők az adott prozódiai egységen a modern francia beszédre jellemző, 2–3 szótagonkénti dallamhangsúlyt.<sup>5</sup> Ennek értelmében, a kiválasztott öt szótagú értelmi és szintaktikai egységen – *n'a/vait/ ja/mais/ eu* [soha nem volt] – az elvárás a 3:2 vagy a 2:3 szótagfelosztás lenne. A francia olvasatok 89%-ában az első felosztást találjuk.

"Monsieur Seguin n'avait jamais eu de bonheur avec ses chèvres."  
(Seguin úrnak soha nem volt szerencséje a kecskéivel.)

1. francia beszélő
2. magyar beszélő nem utánzó kiejtésében
3. magyar beszélő utánzó kiejtésében
4. magyar mondat



11. ábra

A 11. ábrán látható 1. melódiagörbe ezt a sztenderd ejtismódot szemlélteti, amely nagymértékben "elüt" ugyanennek a mondatnak a 100%-osan nem francia kiejtésétől (2. melódiáiv). Ez utóbbit egy nem franciatanár magyar beszélő ejtette ki, aki aztán – saját elhatározásából – a franciákat utánzó formában megismételte első olvasatát. Ennek az utánzó olvasatnak az első mondatát a 3. dallamgörbe szemlélteti, míg a 4. dallamív ugyanezen beszélő magyar olvasatából kivágott

<sup>5</sup> A továbbiakban az egyszerűség kedvéért használt "dallamhangsúly" kifejezést a legújabb kutatások által szorgalmazott prominencia v. "proéminence" jelentésben használjuk.



első mondatot jellemzi. Feltűnő a 2., nem utánzó francia ejtés és a 4., magyar melódiaszerkezet hasonlósága. A 2., nem utánzó és a 3. utánzó kiejtést összevetve a dallamív "hullámossága" és a mondat összsidótartamának rövidülése tűnik fel.

Ez tehát azt jelentheti, amit később az illető saját szavaival is megfogalmazott: számára a francia beszéd gyorsabbnak, és főleg "dallamosabbnak" tűnik, mint a magyar.

Ugyanitt az 1. és 3. dallamív összevetése viszont azt mutatja, hogy az "utánzás" egy ponton nem sikerült, mégpedig éppen azon az ötszótagú dallamegységen, ahol az illető a várt 3:2 vagy 2:3 hangsúlyelosztás helyett 4:1 eloszlást produkált. Ez nemcsak a hangsúlyeloszlás szabályainak, de az adott mondat szintaktikai szerkezetének is ellentmond, mivel így az ötödik szótagra eső hangsúly éppen az *avoir* ige participiumára esik (*eu*), amely soha nem lehet hangsúlyos a franciában.

## Kísérlet

Ez a megfigyelés adta az ötletet a hangsúlyeloszlásnak egy rövid mondaton belüli vizsgálatára. Ez esetben, az egyöntetűség kedvéért a nyolc tesztelt mondatból, az alaphang kivételével, minden információt kiszűrtünk. Így a hallgatók csak a mondat dallamát hallották, a szavakat nem. Egyenkénti, fülhallgatóval történő meghallgatás után a 25 megkérdezett francia hallgatónak feleletválasztós tesztlapra kellett megjelölnie, hogy a beszélőnek "van-e akcentusa", "nincs akcentusa" vagy döntésük "bizonytalan".

A kísérlet sematizált eredményét a 12. ábra szemlélteti. E szerint a 8 szótagos francia mondaton (*et là-haut le loup les mangeait* [és ott felfalta őket a farkas]) a hangsúlyeloszlást a franciák akkor találták megfelelőnek, ha az a 8 szótagú értelmi egységen 3:2:2:1, 3:2:3 vagy 3:2:1:2 formában valósult meg, ill. csakis lexikai vagy értelmi egység végére került hangsúly. (Ld. az 1. 2. és 3. dallamívek.)

Ezzel szemben "akcentust" észleltek azokban az esetekben, amikor pl. a mondatban 2:3:3 szótag eloszlásban elhelyezkedő hangsúly a deiktikus funkciókat el-látó *là* mutatónévmásra, tehát nem lexikai egységre esett (5. dallamív). Egy másik esetben (4. dallamív) hangsúlyhiba történt, mert a mondat végét záró - *aient* szótagnak, kijelentő mondat esetén, csak az időtartama nyúlhat meg (Vaissière 1991), értelemmegkülönböztető szerepe miatt dallamhangsúlyt nem kaphat. Az akcentusosnak ítélt ejtés módok közül két esetben (6. és 7. dallamív) pedig ritmikai okok mutathatók ki, ugyanis a három részre tagolt – a 7. dallamíven még glottális zárral is elválasztott – dallamegységet a franciák "darabosnak", "töredezettnek" találták.

A "bizonytalan" kategóriába sorolt egyetlen ejtés módból csak a többitől eltérő, utolsó öt szótagot emeltük ki. Ez a dallamív arra enged következtetni, hogy

két hangsúly egymás mellé kerülése (3:) 2:1:3 eloszlásban esetleg zavaró lehet. Felmerül azonban a kérdés, hogy az ugyanazokon a szótagokon hasonló hangsúlyeloszlást mutató, de a "nincs akcentusa" kategóriába sorolt 3. dallamtípus miért minősül mégis elfogadhatónak?

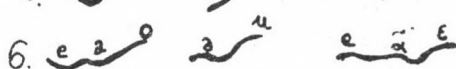
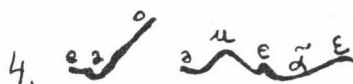
2. "et là haut le loup les mangeait" - /és ott felfalta őket a farkas/

1. csak a melódia vizsgálata után

"nincs akcentusa"

"bizonytalan"

"akcentusa van"



12. ábra

Egy kézenfekvő, de még ellenőrzésre váró magyarázat az lehet, hogy ez utóbbi esetben a két egymás mellé kerülő szótagot dallamesés, ún. glissando választja el egymástól, ami az észlelet szempontjából esetleg megkülönböztető jelentőséggel bír.



A francia nyelvben észlelhető magyar akcentus arányának és okainak fent részletezett vizsgálata alapján elmondható, hogy az 1. és 2. kísérletben kiválasztott két szó, valamint a 3. kísérletben elemzett dallamegységek magyar ejtéseinek túlnyomó többségében, nagy arányban észleltek a megkérdezett francia hallgatók akcentust. Az okok mindhárom esetben anyanyelvi interferenciákra, ill. több nyelvi szint (pl. szintaxis és fonetika) együttes kezelésében mutatkozó nehézségekre vezethetők vissza.

## Függelék

1. "Monsieur Seguin n'avait jamais eu de bonheur avec ses chèvres. Il les perdait toutes de la même façon: un beau matin, elles cassaient leurs cordes, s'en allaient dans les montagnes et là-haut le loup les mangeait. Ni les caresses de leur maître, ni la peur du loup, rien ne les retenait. C'était, paraît-il, des chèvres indépendantes, voulant à tout prix le grand air et la liberté." (Daudet, 1985.)
2. "Seguin úrnak soha nem volt szerencséje a kecskéivel. Mindegyiket ugyanúgy vesztette el: egy szép napon eltépték kötelüket, felszöktek a hegyekbe, és ott felfalta őket a farkas. Sem a gazdájuk simogatása, sem a farkastól való félelem nem tarthatta vissza őket. Úgy hírlík, ezek független kecskék voltak, akik mindenáron szabad levegőre és önállóságra vágytak." (Szöveghű fordításban.)

## Irodalom

- Bakó Elemér: Adalékok a magyar nyelv orrhangú magánhangzóinak kérdéséhez. In: Hangtani tanulmányok. Dolgozatok a Debreceni Tisza István Tudományegyetem magyar nyelvészeti szemináriumából 3. 1937.
- Bannert, R. – Engstrand, O. – Eriksson, H. – Nordstrand, L.: Phonetics and identity. Reports from Uppsala University Dept. of Linguistics n<sup>o</sup>8, 1982, 22–58.
- Calliope: La parole et son traitement automatique. Paris-Milan-Barcelone-Mexico 1989.
- Carton, F.: Introduction à la phonétique du français. Bordas 1991.
- Castelli, E. – Perrier, P. – Badin, P.: Acoustic considerations upon the low nasal formant based on nasopharyngeal tract transfer function measurements. Eurospeech, vol.1., 1989, 412–415.
- Daudet, A.: Les chèvres de Monsieur Seguin, in: Les lettres de mon moulin, Paris 1985.
- Földi, É.: On the nasalization of the Hungarian vowels. Magyar Fonetikai Füzetek 19. 1988, 55–71.
- Herman, J.: Phonétique et phonologie du français contemporain. Budapest 1984.
- Kassai Ilona: A francia és a magyar magánhangzók képzésének kontrasztív vizsgálata. NyK LXXIII, 1971, 441–445.

- Kassai, I.: Etudes contrastives sur le français et le hongrois. Studia Romanica, Kossuth Lajos Tudományegyetem, Debrecen 1974.
- Kassai Ilona: Pszicho-szociolingvisztikai megjegyzések az akcentusról. ÁNyT XVIII. (megjelenés alatt).
- Maeda, S: The role of the sinus cavities in the production of nasal vowels. Proceedings of the ICASSP, vol. 2. 1982, 911–914.
- Olaszy Gábor: Számítógépes beszédelőállítás. Műszaki Könyvkiadó, Budapest, 1989.
- Vaissière, J.: Rhythm, accentuation and final lengthening in French. In: J. Sundberg – L. Nord – R. R. Carlson (eds), Music, Language, Speech and Brain. London 1991, 108–120.



# A SZÖVEGFONETIKAI KUTATÁSOK GONDJAIRÓL

Wacha Imre

ELTE Bölcsészettudományi Kar, Budapest

1. Amikor 1970 körül Szende Tamással kialakítottuk a "szövegfonetika" elnevezést, s 1973-tól kezdve, világviszonylatban is elsőként megjelentek szövegfonetikát célzó cikkeim (Wacha 1973, 1974; vö.: Szabó 1988, 11), azt hittem, **szövegfonetikát** művelek. Később rájöttem, hogy amit közreadtam, csak **mondatlánc-fonetika** volt, és a főbb kommunikációs szituációkban elhangzó beszéd hangzásvilágának (akusztikumának) általános jellemzése, de nem "a" "szöveg", fonetikai jellemzőinek (fonetikai szerkezetének) leírása, rendszerezése. Mentségül szolgálhat, hogy Magyarországon csak ekkoriban kezdődtek meg a szövegtani kutatások. Meggyőződésem: kiváló eredményeik ellenére, hosszú ideig ezek is csak mondatlánc-kutatások voltak.

A magyarázat az lehet, hogy valójában még ma sem sikerült megnyugtatóan megegyezni abban, mi az, amit a *szöveg* terminussal nevezünk meg. Így bizonytalan az is, mi is tartozik a *szövegfonetika* körébe. Mindezek ellenére mind a szövegtani, mind a szövegfonetikai kutatások jelentős eredményeket tudnak felmutatni. (Hosszú ideig a mondatnak sem volt megnyugtatóan egyértelmű meghatározása. Azért a mondatnani és a mondatfonetikai kutatások eredményesen folytak.)

2. Természetesen a **szövegnek** több, szövegtani és szövegfonetikai szempontból is jó, elfogadható meghatározása van. A magyar szakirodalomban általában Deme László definícióját szokás elfogadni: "A szöveg nyelvi formába öntött objektívációja az egyéni pszichikai tartalom egy részletének, olyan terjedelemben és megformáltságban, amely elegendő ahhoz, hogy adott helyzetben, a kifejezés és/vagy tájékoztatás és/vagy befolyásolás feladatát ellátva, a teljesség és lezártság érzetét is felkeltse" (Deme 1979b, 59; vö. még: Deme 1980, 11, 1983, 1988, 1988/89. 44)<sup>1</sup>, kiegészítve a hangzó szövegre Bolla Kálmán meghatározásával: "a szöveg a beszédműnek a megnyilatkozásnál nagyobb tematikus egysége, amelynek rá jellemző szemantikai, grammatikai és fonetikai felépítése (megszerkesztettsége) van, s a beszédhelyzet függvényében nyeri el és tölti be funkcióját" (Bolla 1989, 55)<sup>2</sup>.

E meghatározások azonban a *szöveg* szakkifejezésnek többféle értelmezését is lehetővé teszik.<sup>3</sup> Így nyugodtan állíthatjuk, hogy a szöveg fogalmának nincsen egységesen elfogadott meghatározása. Ezért a kutatók többsége azt az általa vizsgált korpuszt tekinti szövegnek, melyet valamiért egységnek/egésznek fogad el. (Pl. egy teljes irodalmi vagy írásművet vagy annak egy többé-kevésbé kerek,

önálló egészként is elfogadható részletét, egy – egyetlen – gondolati egységet tartalmazó, általában nem túl rövid terjedelmű beszédművet, olykor csak egy-egy bekezdésnyi gondolatsort, de olykor néhány, sőt egyetlen mondatból álló közlésegeset is.) Előfordul az is, hogy vizsgált korpusz határait mechanikusan húzzák meg, olykor a szövegegészet is megkurtítják.

Nem véletlen, hogy a szövegtan kutatóinak jelentős része **csak az írott szöveg** és főleg a jól megszerkesztett folyamatos (különböző típusú és főleg művészi-irodalmi) szöveg szövegtani vizsgálatával foglalkozik. A spontán szöveggel, főleg a párbeszéddel nem nagyon tud mit kezdeni.<sup>4</sup> Nem véletlenül teszi fel tehát a kérdést Kocsány Piroska, hogy a szövegnyelvészet valójában nem szövegtypológiai nyelvészet-e (Kocsány 1989).

Az sem véletlen, hogy a hazai szövegfonetikai (sőt a fonetikai) kutatások is (ide értve a helyes szövegmondás kérdéseivel foglalkozó cikkeket is) – a Kazinczy-versenyek, a rádiós, televíziós és a színpadi beszéd problematikájához kapcsolódva – az **írott szöveg** reprodukálásának, interpretálásának kérdéseire figyeltek fel, annak akusztikus jellemzőit, megszólaltatásának törvényszerűségeit vizsgálták elsősorban; bizonyos értelemben éppen a nyelvi helyesség, tartalmi-logikai adekvátság és a kiejtési vagy beszédnorma megállapításának érdekében, már a múlt század közepe óta. (Vö. Wacha 1991b. Vö. még: Deme 1964, 1965, 1973, 1975, 1976, 1979a, 1988, 1992; Bolla 1989; Elekfi 1968, 1979, 1980; Fónagy 1974; Fónagy – Magdics 1967; Wacha 1973, 1991a-b, 1992a-b; Z. Szabó – Wacha 1978, 1991.) Fónagy Iván és Magdics Klára könyvüknek (1967) jelentős részében szintén a szöveginterpretálás alapján elemzik a "szöveg"-részlet intonációs formáit. (Megfigyeléseik, a megállapított nyelvi törvényszerűségek a szövegfonetikai kutatások előzményének tekinthetők.)

Hogy a kutatások nem nagyon nyúltak a spontánul beszélt szöveghez, annak több oka is lehet. Pl. **a)** a lejegyzéskor, vizsgálatkor roppant nehéz elkülöníteni a szöveg meg- és beszerkesztett "alapegységeit", a mondatoknak megfelelő megnyilatkozásokat. (Vö. Balogh 1972, 1987; Németh T. 1988, 1989; Szabó 1983; Szende 1976, 158; Wacha 1988, 1991b-c.) **b)** Nehéz a szövegegység, a szövegegész határainak megállapítása, s így a vizsgálandó korpusz határainak megvonaása, szemben a jól megszerkesztett írott szövegekkel. (A szövegfonetikai kutatások is ilyenek hangosított, interpretált változatát elemzik. Vö.: Bolla 1991, 1992; Bolla – Földi 1992, valamint az EFF. 1–2. számának szövegelemző tanulmányait.) **c)** A spontán szövegek vizsgálatakor még sok, a vizsgálatokat megnehezítő tényezővel kell számolnunk. Pl. mikor lépünk át a mondatfonetikai vizsgálódásokból a szövegfonetikába, mi legyen a mondattani szempontból henyé szövegrészletekkel, nem foghatók meg a nonverbális elemek, stb.

**3.** Ne foglalkozunk tovább a szöveg meghatározásának kérdéseivel! A **szövegfonetikai kutatások szempontjából pillanatnyilag ez még közömbös kérdés.** Ugyanis mielőtt a **szöveg** fonetikai szerkezetének vizsgálatához kezdenénk,



– azt hiszem – előbb le kell választani a hangzásformáról olyan elemeket, melyek talán nem is a szöveg fonetikai szerkezetének jellemzői, mert **elsősorban** arra a **kommunikációs szituációra, kommunikációs típusra** jellemzőek, amelyben a szöveg létrejön és elhangzik (noha ennek a szövegforma létrehozásában–kialakításában és a hangzásforma alakulásában is determináns szerepe van). (Vö. pl.: Deme 1973, 1991, 1992; Wacha 1973, 1985, 1991a, 1992a-b, d.)

A szöveg ugyanis, a következő főbb szituációkban jöhet létre és hangozhat el:

(**aa**) közvetlen, totális – azaz verbális és nonverbális – kommunikáció eredményeként spontán szövegalkotással (beszédműként); azután

(**ab**) korlátozott kódú – azaz csak írásos – szövegalkotás eredményeként (írásmű); illetőleg

(**ac**) írásos/írott alapról induló, de nem kötött szövegalkotású kommunikáció eredményeként (interpretáció).

Megjelenhet, betöltheti kommunikációs funkcióját (**aba**) írott formában (írásmű) és (**aaa** és **acb**) akusztikus (hallható) formában (beszédmű).

Ez utóbbiban lehet (**aaa**) spontán élőszó és (**acb**) az írott szöveg hangosítása: (**acba**) felolvasás, (**acbb**) értő/értető tolmácsolás, (**acbc**) interpretálás.

Hagyjuk most figyelmen kívül ezt és a spontán beszéd(szöveg) szövegfonetikai szerkezetének vizsgálatát is. (Vö. azonban Art 1991; Szabó 1883; és a Beszélt nyelvi tanulmányok című kötet tanulmányaival [Kontra 1988].)

A szövegfonetika gondjainak felvázolására elegendő, ha az írásos fogantatású, írásos alapról induló és elhangzó beszéd hangzásának problematikájáról szönlünk.

Tudjuk, vizsgálatok, cikkek bizonyítják, hogy (**aa**) a spontán élőszó, (**ab**) az olvasónak szánt írásmű és (**ac**) az elmondásra, interpretálásra szánt (írott) szöveg és nyelvi formája, következésképpen akusztikus formája is, meglehetősen különbözik egymástól.<sup>5</sup> (Vö.: Deme 1974; Wacha 1973, 1985, 1991a, 1992a-b, d.)

Tekintsünk el attól is, hogy az írott szöveg háromféleképpen szólalhat meg: **a**) felolvasásként, **b**) értő--értető tolmácsolásként és **c**) interpretálásként, azaz művészi értékű, élőszó"szerű" tolmácsolásként. (Vö. pl.: Bolla 1989, 55; Deme 1991, 1992; Subosits 1982, 14; Wacha 1973, 1974, 1985, 1991, 1992b, d.).

Tekintsünk el attól is, hogy a megszólaltató papírról olvassa-e a szöveget vagy betanulva mondja el, attól is, hogy saját írott szövegét szólaltatja-e meg a beszélő.

Mindezek **lényeges** hangzásbeli különbségeket adhatnak azonos szöveg esetén is! (Nem részletezem, eléggé sokat írtam erről. Vö. mégis pl.: Wacha 1973, 1974, 1985, 1992b, d.) Tekintsünk el mindettől!

4. Még így is gondot jelent, hogy a létrejött akusztikus formából **mi képezi a szövegnek a szövegségéből adódó fonetikai szerkezetét.**

**4.1.** A rádiót hallgatva tapasztaljuk, hogy lényeges különbség van a hírek, tudósítások, kommentárok, jegyzetek hangzása között: mindegyik **műfajnak** megvannak a sajátos, **nyelvi és akusztikus jellemzői**, bárkinek a szájából hangzanak is fel ezek a szövegek. Lehet mondani persze, hogy ezek a szöveg műfajából (nyelvi formájából) adódnak és a hangzásforma ezeknek a szövegeknek az akusztikus szerkezetét tükrözi. Úgy tűnik, itt teljesen jogos a szöveg fonetikai szerkezetének vizsgálata. (Vö. erre: Deme, Wacha, idézett munkáit.)<sup>6</sup>

**4.2.** Mégis kétségeink támadhatnak, hiszen az elmondottakhoz hasonló "stílusjellemzőket" (?) megfigyelhettünk a külföldi: a (kelet)német, a bolgár, a szerb, az orosz rádiók adásait hallgatva. Hasonló jelenségeket tapasztalhatunk az angol, francia adásokban is. Pedig e nyelvek szerkezete más, mint a magyaré. Így joggal következtethetünk arra is, hogy az azonos **kommunikációs szituációban**: egyirányú rádiós, azaz kevert kommunikációs helyzetben (vö. Wacha 1985, I, 49–66, 1992c, 23–7), értő/értető szándékú felolvasással, más-más személyek, eltérő egyéniségek szájából felhangzó szöveg hasonló **akusztikus jellemzőit** nemcsak a szövegek műfaji és stílussajátosságai alakították ki. A **közléshelyzet** és a **közlésszándék** is formálta az "uniformizált(nak ható)" "bemondói stílust". (A közlésszándék formáló – "jelentéstorzító" – hatására l. az Egy gondolat bánt engemet Petőfi-versnek Antal Imre által megszólaltatott akusztikus paródiáját.)

**4.3.** Ugyancsak erősítik az imént elmondottakat azok a kísérletek, amikor irodalmi szövegeket "beszélgetésszerűen", az élőszó hangvételével próbáltunk megszólaltatni. A szövegek ellenálltak; irodalmi hangvételre "kényszerítették" a szövegmondót. (Annak idején a kiváló előadóművész Berek Katinak is csak egy verset sikerült igazán "beszélgetésszerűen" megszólaltatnia: Petőfi Anyám tyúkjá című versét.)

Következtetésünket gyengíthetik is, erősíthetik is más megfigyelések.

**4.4.** A pódiumon vagy a rádióban, tévében megszólaló irodalmi szövegeknek is sajátosan egységes hangzásuk és stílusuk, "irodalmi hangvételük" van, függetlenül szerzőjüktől, műfajuktól és keletkezési koruktól. S attól is, ki szólaltatja meg a művet. A Kazinczy-versenyek kapcsán "pódiumstílus"-nak, "generálakusztikum"-nak neveztem azt a hangzást, mely eluralkodik a szöveg felett, bármilyen is a szöveg szerkezete, nyelvezete. (Wacha 1991a, 87–8, 1992a, 66–7.)

**5.** Úgy tűnik tehát, hogy a szöveg "hangosításakor" fölébe nő a szövegben magában (mondataiban, mondatláncában) is rejlő latens zenének, a **szöveg** akusztikus szerkezetének egy olyan (a nyelvi közösségtől is elvárt) akusztikus forma (pl. emelt hangfekvés, kiénekelte dallamok stb.), ami a **közléshelyzetre** és a **közlésmódra** jellemző. "Megvalósítását" a beszélők szinte kötelezőnek érzik. Ez akkor is tapasztalható – s nemcsak a magyar nyelvre jellemzően –, ha a szer-



ző, beszélő adja elő, mondja el saját művét, saját szövegét. (Vö.: Wacha 1991a, 1992a; Z. Szabó – Wacha 1991. *passim*.) Gondoljunk csak a lemezlovas- (disc-jockey) stílusra vagy az angliai lóversenyközvetítések stílusára!

6. Mindezt csak részben módosítja az, hogy a **nagy előadói egyéniségeknek** (pl. Ascher Oszkár, Berek Kati, Csernus Mariann, Gábor Miklós, Gáti József, Horváth Ferenc, Jancsó Adrienne, Keres Emil, Latinovits Zoltán, Mensáros László, Ódry Árpád, Pálos György, Somlay Artúr, Tímár József) megvolt és megvan a maguk sajátos stílusuk, mely beleillik a pódium megkövetelte hangzásba, de ugyanakkor önmagukon belül is egységes stílusra mutat, tolmácsolásukban némileg egységes hangzásúvá teszi a különböző szerzőktől származó különböző mondanivalójú és műfajú irodalmi alkotásokat is.

7.1. De hogyan értékelhetjük szövegfonetikai szempontból, amikor ugyanazon irodalmi műnek jelentős mértékben megváltozik a hangzásvilága, a korstílusoknak, a rendezői, színészi egyéniségeknek megfelelően. Tapasztalhatjuk még akkor is, ha ugyanazon műnek azonos korban készült színpadi, rádiójáték és filmbeli feldolgozásának hangzását vetjük össze. (Az ember tragédiája jó példa minderre, hiszen maradtak fenn felvételek különféle színrevitelekről. Vö.: Wacha 1989, 1990.)

7.2. Hasonlókat tapasztalhatunk akkor is, ha a versmondást figyeljük meg. A Varietas delectat gyűjtőnévvel megjelentetett verses lemezek felvételeinek hangvételén is átsüt az irodalmi vagy "pódium-stílus", a korstílus, de átsüt rajtuk az interpretáló egyéni stílusa is. (Vö.: Molnár – Surányi 1981.) Hol van, hogyan található meg a szövegben latensen meglévő hangzás, a **szöveg** akusztikus szerkezete?

7.3. Ilyesféle gondok merülhetnek fel (az előre megírt) közéleti, a szónoki beszédek hallgatva is. Ha közülük a prédikációkat vizsgáljuk, tapasztaljuk, van valami hangzásukban, ami csak rájuk jellemző. Ezen belül azonban olyan eltéréseket is találhatunk, melyekkel (pl. a beszéddallam és beszédrítmus\* által) felismerhetjük, hogy mely felekezetek papjának, katolikus, református, evangélikus, zsidó papnak az igehirdetését halljuk. A katolikuson belül még azt is elkülöníthetjük, hogy pl. bencés, szalézi szerzetes vagy világi pap prédikál-e.<sup>7</sup>

8. Mégis jogos a szöveg akusztikus szerkezetét vizsgálni. Az előbb felsorolt **eltérő hangzásformákon belül** is találunk olyan akusztikus megoldásokat, melyek olykor azonos, vagy hasonló, vagy legalább tendenciájukban hasonló/meg-

\* Beszédrítmuson itt nem az egyes hangok időtartambeli különbségeiből adódó ritmust – a magam nem pontos kifejezésével: artikulációs ritmust – értem, hanem elsősorban a beszédnek azt a lüktetését, amely részben – a hagyományos értelemben vett – hangsúlyozásból, részben a hangsúlyok és a hanglejtés által összefogott különböző hosszúságú nyelvi egységek gyorsabb-lassabb kimondásának lüktetéséből adódik. Egy olyanfajta lüktetés ez, mint a magyaros ritmus a verselésben.



felelő akusztikus formákat eredményeznek (pl. hangsúlyozásban, dallamformában, tagolásban, ritmusváltásban). Ezeket a variánsokat közös alapformára lehet visszavezetni oly módon, mint Varga László tette a magyar beszéd prozodémáinak bemutatásakor (1989). Ezek már olyan elemeket mutatnak, melyek már nem a szituációra, nem a megszólaltató személyiségére, hanem a műfajra, műnemre, sőt a megszólaltatott szövegre jellemzők.

9. Az eddigiekből arra következtethetünk, hogy a szövegtani kutatásokhoz hasonlóan (vö.: Kocsány 1989), a szövegfonetikai kutatásokban is először egy-egy szöveg(típus)nak kell feltárnunk az akusztikai szerkezetét. Ezt úgy kaphatjuk meg, hogy először egy-egy szöveget több személlyel szólaltatunk meg, s leírjuk, elemezzük az így kapott különböző hangzó utánezeteket. (Módszerére: Bolla 1991, 1992; Bolla – Földi 1992.) Összevetésük után kiválasztjuk a közös jellemzőket. A variánsok elemzése után megkaphatjuk a **szövegtípusok közös akusztikus** – főleg szupraszegmentális szintű – **jellemzőit**. Persze közben ügyelnünk kell a szövegnek (megfigyeléseink alapján) kiemelt, kritikus (szövegtani) jellemzőire (bevezető, lezáró rész, témaváltó, -ismétlő részek, logikai egységek határai stb.), struktúrajegyeire, és ügyelnünk kell a mondat- és szövegfonetikai eszközöknek (eddig felismert) funkcióira (kapcsolás-tagolás, lényegkiemelés, kommunikációs szervezés stb.), funkció sajátosságaira. (Vö. pl. Bolla 1989, 56.)

A szöveg fonetikai eszközei természetesen ugyanolyan szupraszegmentális eszközök, mint a mondatfonetikaiak, de már nem a mondaton belülre utalnak, hanem a mondaton túlra, a mondatláncra, a szövegegészre. Használatukat egyrészt a mondatlánc és a szövegegész tartalmi, logikai, grammatikai stb. felépítése, információs súlyviszonyai, összefüggései determinálják (előbb a szöveg, a kontextus "nyelvi" szintjén), nem pedig az "önálló" mondat információs és szintaktikai relációi. A mondatláncbeli beszerkesztettség, az információs súlyviszonyokra a mondatláncból kiszakított szövegmondat is utal, pl. megszerkesztettsége, szintaktikai szerkezete, a mondatrészek sor- és szórendje, a kataforikus, az anaforikus és a deiktikus elemek, valamint az ezeket tükröző intonáció (valamennyi szupraszegmentális eszköz) által (vö. Deme 1979a, 1991, 1992; Wacha 1992a).

A szövegfonetikai eszközök használatát megszabja az (el- és felhangzó) szöveg intertextuális helyzete, a beszédhelyzetbe, az előzményekbe való be- vagy be nem épített volta, a partnervizonyok, a partnerek előismeretei stb. (Vö.: Deme 1991, 1992; Wacha 1991a, 1992a.)

#### **A mondatfonetikai eszközök funkciói a szövegben:**

A **szövegmondat** síkján a **mondatfonetikai** eszközök szerepének ellátása:

**a)** a mondat aktuális tagolása, a felvezető, újságoló, mondathangsúlyos részek elkülönítése;

**b)** az információs súlyviszonyok tükrözése a szövegmondatokban;



c) az egyes szövegmondatok szintaktikai szerkezetének érzékeltetése (a szintagmák összefogása, tagolása a mondategységek, mondatrészek elkülönítése–kapcsolása;

d) a mondat jelentésének (kijelentő, kérdő stb. voltának) érzékeltetése;

e) a szövegmondat bekapcsolása a mondatláncba, pl. a befejezetlenség, folytatás szándékának kifejezése, az elhangzó (mondat)szerkezetre következő (mondat)szerkezet jellegének (előre)jelzése.

A **szövegfonetikai eszközök funkciója** többsikú:

a) A **szöveg** (kontextus) síkján a **mondatláncban szövegfonetikai** funkcióban:

aa) az információs súlyviszonyok érzékeltetése a mondatláncban;

ab) a szövegmondatoknak kapcsolása a mondatlánc előző és következő láncszeméhez (előre- és visszautalás), s ha kell eltagolása attól;

b) **Kontextuális** szerep a **szövegegészben**:

ba) a szövegmondatok egymáshoz viszonyított információs súlyának érzékeltetése a szövegegészben;

bb) a tételhangsúlyos és a szöveghangsúlyos részek kiemelése;

bc) a távolra szakadt részeknek (pl. tételmondatoknak stb.) akusztikus összekapcsolása, különösen, ha a nyelvi forma nem eléggé pontos;

bd) a szövegen belüli attitűdváltozások jelzése;

c) **Intertextuális** szerep: a szöveg, szövegmondatok bekapcsolása az intertextusba és a kontextusba.

d) **Kontextuális és intertextuális** funkció: a szöveg szerkezetének, építkezésének érzékeltetése, a szövegszerkezet tagolása, ábrázolása.

e) A szöveg **műfajának, közlésmódjának** érzékeltetése.

f) **Intertextuális funkció**: az attitűdök és attitűdváltások érzékeltetése.

Természetesen további funkciókról is szólhatnánk még, de azok már talán a következő, a nagyobb "egység", a "beszédegész" vagy talán "közlésegész", a "beszéd", a "kommunikációs folyamat" (vö.: Deme 1979, 58–9) illetékességi körébe tartoznak.

10. Ahhoz, hogy a **szövegnek** a fonetikai szerkezetét, s a szövegnek szövegségét megadó fonetikai szerkezetet megkaphassuk, először le kell választani a hangzásformáról azt, ami mintegy külsődleges elem, mert független a tartalomtól, a szövegformától: a közléshelyzetet, közlésnemet (közlésformát, szituációt) és a szövegtípussal kapcsolatos közösségi hangzási, hangvételi elvárásokat tükrözi. Ezek után következhet a műnemre, a műfajra, majd a megszólaltató személyre jellemző hangzásformáknak a leválasztása, s csak ezután következhet annak megvizsgálása, mi az, ami most már magára a szövegre, a szöveg (s "a" "szöveg") szerkezetére, logikai, nyelvi felépítésére jellemző, nyelvi formájának (a szövegmondatok szintaktikai szerkezetének, a mondatok közötti kohézióinak),

a tartalmi és nyelvi szempontból újak, fontosnak, magától értetődőnek a kifejezője. Ilyen tekintetből különösképpen meg kell vizsgálni a szövegnek az ún. kritikus pontjait: a szöveg indítását (ez adja meg a "hangvételt", ez kapcsolja be a szöveget az intertextusba), a kisebb és nagyobb logikai egységek határát (miképpen tagolja az akusztikus szerkezet itt a szöveget, és miként kapcsolja össze az egységeket), végezetül pedig a szöveg zárlatát. Ezen belül a bekezdésnyi (új témát, témarészletet indító–tartalmazó) egységeknek, majd a mondatoknak a szupraszegmentális szerkezetét (a kiemeléseket; a szerkezeti tagolását a mondatnak; a – távolra szakadt vagy egymás után következő, de valamiképpen kapcsolódó – mondatok közötti kohéziót, pl. mondatzárlatokat vagy a mondatvégek "átértékelését" stb.).

A munkálatok ilyenén elvégzésével megkapjuk előbb a **szöveg fonetikai szerkezetét**, majd a **szövegfonetika szerkezetét**.

Természetesen más kérdéseket vet fel, másfajta vizsgálatokat (is) igényel a kötetlen, spontán beszéd kutatása. Előbb talán az előadásszerű megnyilatkozásoké, majd a dialógusoké. E téren a legégetőbb kérdés ugyanis, hogy meddig tart a szöveg, hol van a szövegegység vagy a szövegegész határa.

**11.** Az elmondottakon kívül további gondok, nehézségek is állnak a szövegfonetikai kutatás előtt. A szövegfonetikai kutatások munka-, idő-, műszer-, energia-, azaz pénzigényesek. Ráadásul sok munkával viszonylag kevés látványos eredményt hoznak. És e téren úgy állunk, mint az egyszeri plébános a püspök látogatásakor: harangoztatott volna szegény feje, de a templomnak nem volt harangja.

### Jegyzetek

- <sup>1</sup> Deme különben A beszéd, a nyelv, a szöveg című cikkében (Deme 1988/9, 47–54) és másutt is (pl.: Deme 1974, 12–20, 116–21), többször is kimondja, hogy a megszerkesztettség nem feltétlen jellemzője a szövegnek.
- <sup>2</sup> E meghatározásban gondot jelenthet a "megnyilatkozásnál nagyobb tematikus egység", mert nem tudni, hol vannak ennek alsó és felső határai. (Erre nézve vö. még Deme 1979a, 58–9, 64–5; Szende 1976, 158. A szöveggel kapcsolatos meghatározásokra vö. még Balázs 1985; Petőfi S. J. 1990, 1991; Vass 1990 stb.) A vizsgálódásokban egyébként is nagy gondot jelent az érintkezési pontok megtalálása. Így például, hogy a szövegtani vizsgálódások szempontjai nem mindig, sőt rendszerint nem felelnek meg a szigorúbb értelemben vett fonetikai vizsgálódások szempontjainak. Érdemes ilyen alapon megfigyelni Szende Tamás összevetését, melyben az információelméleti, a grammatikai és a fonetikai vizsgálódások közös pontjait keresi (Szende 1976, 141–61; vö.: Kassai 1981, 63–6.). Ezt most megkísérlem kibővíteni – gondolom nem hibátlanul – a szövegtan szempontjaira is figyelve. A [...] jeles részek tőlem származnak. Szende nyomán a következőket állíthatjuk:  
Ha a verbális produktumot (akár írott, akár hangosított, akár spontán megfogalmazású beszélt szöveg legyen is) – a rendszerezés kényszeréből és szándékából – megpróbáljuk szerkezeti–szerveződési egységeire bontani, alulról (a legkisebb nyelvi elem irányából) felfelé építkezve maguktól adódnak a következő egységek a verbális megnyilatkozás tagolására. **Információelméleti** szempontok figyelembevételével a következő egységekkel számolhatunk: hang – szó



– mondat – [gondolati-logikai (rész)egység (részegész)] – beszédmű [(beszédegész) (kontextus)]. A **fonetikai** szempontú rendszerezés egységei: hang(változat) – hangkapcsolat – szótag – fonemikus frázis – szólam – [(?.)] – beszédmű mint folyamat. **Grammatikai** szempontból a következő szerveződési egységeket kell figyelembe vennünk: fonéma – morféma – szó – [(?szintagma?)] – mondat – [nyelvi logikai egység (bekezdés)] – beszédmű.

Táblázatosan:

Információelméleti szempont	Fonetikai szempont	Grammatikai szempont
hang	hang(változat)	fonéma
----	hangkapcsolat	-----
----	szótag	-----
----		morféma
szó	-----	szó
[szószerkezet]	fonemikus frázis	[(?szintagma?)]
----	szólam	[(?szintagma?)]
mondat	(?nyilatkozat?)	mondat
[gondolati logikai (rész)egység]	[(?.)]	[bekezdés]
beszédmű (kontextus)	beszédmű (mint folyamat)	beszédmű (kontextus)

3 Pl.: **a)** Minden, ami (valamilyen kommunikációs vagy kommunikatív vagy informatív célból) elhangzik az első hangtól az utolsóig, illetőleg, mindaz, ami leíratik, az első nagybetűtől az utolsó írásjelig.

**b)** Olyan írott vagy elhangzó verbális szerkesztmény, ami megfelel a **szövegség** kritériumainak (hogy mik ezek, most mellőzzük el) és betölti kommunikációs funkcióját.

**c)** Sima írásmű, annak többé-kevésbé – mechanikusan – körülhatárolt része.

**d)** (Mint leggyakoribb értelmezés:) Irodalmi értékű szöveg, amely egyrészről megfelel a szövegség kritériumainak, másrészről többletként eleve magában hordozza vagy megteremti a kommunikáció négy, illetve öt feltételét: a – közös – valóságról szólást, a közös nyelvűséget, a közös előismereteket, a közös előzményeket, s a közös beszédhelyzetet. (Vö.: Deme 1979, 1980, 15–22). Ezen kívül azonban tartalmi és esztétikai többletet is hordoz: a szövegnek "szó szerinti" (jelentés)síkja mögött és fölött van második, harmadik jelentéssíkja is: az "alapszöveg" asszociációkat ébreszt, jó értelemben vett "aktualizálásra" is alkalmas, azaz a "sztorin" felüli, a mindenkori mának, szóló mondanivalója is van. A szavak hordozta jelentésen túl tehát rejtetten tartalmaz és ezáltal kifejez valami "többletet", általános emberit, mely időállóvá, sőt időn felülivé teszi főleg az értő, de gyakran a laikus számára is. – Kivétel persze számosat találunk. Jó néhány irodalmi mű, költemény, groteszk stb, ugyan nem felel meg a szövegség (de-me) követelményének, mégis szöveg, és mégis jó, irodalmi értékű szöveg. Ezekkel vagyunk a legnagyobb bajban az interpretáláskor, de az értelmezéskor is. Az ilyen szöveget viszont Petőfi S. János gyakran belefoglalja a szöveg fogalomkörébe. Persze bizonyos megkötésekkel.

A **b)** és **d)** csoportba sorolt szöveg rendszerint írott formában, írásműként jelenik meg (előszőr), s hangzó változata többnyire reprodukció vagy interpretáció eredménye. Ilyen értelemben tehát **verbális szövegnek** csak valamilyen tudatos nyelvi szerkesztmény tekinthető, valami olyan manifestációja a gondolkodásnak, mely írásban is meg van fogalmazva és rögzítve van, s ha "megszólal" (lehetőleg) írott alapra épül, sőt hangzó formája ennek akusztikus realizálódása, s mind írott, mind hangzó változata lezárt, kerek egész. – Általában igényesebb irodalmi mű.

Amikor a szövegtani és szövegfonetikai kutatások megindultak, a kutatók többsége csak ezt a lezárt, kerek, írott–irodalmi szöveget tekintette szövegnek, s ezt, illetve hangosított (interpretált) változatát vizsgálta szövegtani és szövegfonetikai szempontból. Csakhogy van olyan szöveg is, mely nem felel meg a szövegség kritériumainak, mégis betölti kommunikatív vagy informatív funkcióját.

És lehet olyan szöveg is, mely ugyan nem tölt be kommunikatív funkciót, nincs is ilyen célja, de mégis kifejez valamit, tájékoztat valamiről, s ezért szövegnek kell tekintenünk. Pl. magában beszélő kisgyerekeknek, felnőttnek vagy akár egy elmebetegnek összefüggéstelen monológja.



Nb. József Attila szabad asszociációinak jegyzéke bizonyára kevésbé felel meg a szöveg retorikai értelemben vett követelményeinek, s a szövegség fogalmának. Azért az mégiscsak szöveg. S van "adekvát" hangzása is!

- 4 Talán éppen ezért nem a legszerencsésebb a *szöveg* szó használata, mert a *szöveg* elem az **írott** szövegre irányítja a figyelmet, s eltereli érdeklődésünket a nem-írott, a spontán beszédről, mely nemegyszer éppen annak van híjával, ami a(z írott) szöveget szöveggé teszi: a tematikus egységnek, a szemantikai és grammatikai megszerkesztettségnek. Elméleti és gyakorlati meg gondolásból talán célszerűbb lenne a *szöveg* helyett, ilyen értelemben, a semlegesebb jelentésű *kommunikátum* szót használni 'tartalmi szempontból vagy a közlésfolyamat befejezettségét tekintve egységet alkotó beszéd- vagy közlésegsz' jelentésben. A *szöveg* terminus technicust elsősorban az **írott** formában megjelenő szövegre és az írásos alapról induló beszédre, a reprodukcióra, interpretációra használnám. Hiszen a szövegtani és szövegfonetikai kutatások főleg azokat a nyelvi és nem nyelvi, nyelven kívüli vagy nyelv feletti elemeket, eszközöket gyűjtik össze és vizsgálják, melyek a közlésfolyamatban egymás után következő megnyilatkozási egységeket (mondatokat, megnyilatkozásokat) összefüggő mondatláncá, megnyilatkozásláncá, egységes szöveggé fűzik és tartják össze. Azt tehát, hogy a megnyilatkozás során mik a szervező eszközök és erők.

Azt is vizsgálják, hogy az egyes közlésmódokban, közléshelyzetekben, közlési műfajokban, műnemekben a közlő a nyelvi eszközökből mit, milyen mértékben és hogyan használ fel mondanivalójának kifejezésére: a közlemény egészében milyen nyelvi eszközöknek, milyen jellegű és mértékű használata jellemzi az egyes közlésmódokat, közléshelyzeteket, műfajokat és műnemeket, stílusokat. A "szövegség" követelményének megfelelő szöveg vizsgálatán belül a kutatók egyrészt a szöveg logikai, szerkezeti építkezését, s ennek nyelvi vetületét vizsgálják (pl. Békési Imre munkásságának jelentős részében), részben pedig a különféle műfajokba sorolható – és főleg írott – szövegeknek szövegszerkezeti, szövegszerkesztési törvényszerűségeit, a szöveg szövegszerkezetének minőségi, mélységi és mennyiségi mutatóit, a szövegmondatoknak mondat szerkezeti és mondat szerkesztési törvényszerűségeit. (Vö. pl. Art Valéria 1991; Balázs 1985; Békési 1986; Bolla 1988; Deme 1961, 1962, 1964, 1965, 1971, 1974, 1976, 1978, 1989; Dezséry – Terestyéni 1976; Elekfi 1968, 1980; Fábricz 1988; Fónagy – Magdics 1963, 1966, 1967; Hajnal Jenő 1991; Kassai 1981; Keszler 1983; Petőfi S. J. 1990, 1991; Szabó 1983; Terestyéni 1987; Vértés O. A. 1987a-b; Wacha 1973, 1974, 1985, 1989, 1991a, 1992b; stb.)

Itt jegyzem meg, hogy amit az irodalmi stilisztika az irodalmi szövegek nyelvében *beszélt nyelvi (fordulat)*-nak minősít, azt – annak ellenére, hogy a beszélt nyelvből került át az írott nyelvhasználatba, az írott stílusba – inkább *társalgási nyelvi (fordulat)*-nak kellene nevezni.

- 5 Éppen az ebből fakadó gondok adják a retorika, a Kazinczy-versenyek, a Beszélni nehéz adások problematikáját, anyagát, s ebből a helyes kiejtéssel – a hangsúlyozás, hanglejtés kérdéseivel –, a színészi beszéd gondjaival foglalkozó nyelvészeti és tudománynpszerűsítő szakirodalom jelentős része. (Cikkekre, bőséges számuk miatt nem hivatkozom.)
- 6 Ezt látszik igazolni az is, hogy bár 1964 óta próbálkozunk vele, kismértékben sikerült csak a rádiós hírmondás akusztikus stílusát megváltoztatnunk. (Vö.: Rádbesz. 1973. passim, Wacha 1985.) Több nyelvész szerző véleménye szerint a legfőbb akadály a rádióhírek megfogalmazása volt. A rádióhírek, kommentárok hangzásában némi változás akkor következett be, amikor szerzők mondat szerkezeteiket az élőszó mondat szerkesztési szokásaihoz igazították.
- 7 Az intonáció alapján talán még a rétori iskolák hatását, tanintézetek kirajzását is nyomon lehet követni, mint a nyelvjárások alapján a népességi migrációt. (Lásd erre Szabó József kutatásait.) A magyar protestáns prédikációs stílusról szólva szokás beszélni debreceni, pápai, pataki stílusról. Ezek állítólag a nyugateurópai protestáns egyetemi városok (Heidelberg, Groningen, Wittenberg és főleg Genf) hatását tükrözik. (Ez bizonyításra vár.) Szathmári István szíves szóbeli közlése, hogy a párizsi (francia nyelvű) református prédikáció hanglejtése és gesztusa őt a debreceni prédikációs stílusra emlékeztették. Magam tapasztalata: 1990-ben a tévé magyar nyelvű misét közvetített Gyulafehérvárról. A prédikáló pap hanglejtése feltűnően emlékeztetett a (pannonhalmi és győri) bencés szerzetesek prédikációjának hangzására. (Gyulafehérváron egykor



volt bencés apátság.) – Az iskolai hatások olykor nagyon erősek. Az intonáció (főleg a hanglejtés) alapján a Kazinczy-versenyek zsűritagjai többnyire felismerték az egyházi iskolába járó diákokat és a teológiai hallgatókat.

### Irodalom

- Art Valéria: Szöveg típus – szövegszerkezet – fordítás. Magisztériumi dolgozat. Újvidék–Szabadka, 1991. Kézirat. 1991.
- Balázs János: A szöveg. Budapest, 1985.
- Balogh Lajos: A nyelvjárási szöveglejegyzés kérdései. Kézirat. 1972.
- Balogh Lajos: A területi nyelvváltozatok fonetikai–fonológiai leírásának problémái. MNy. LXXXIII, 1987, 410–5.
- Békési Imre: A gondolkodás grammatikája. – Szöveg- és mondat szerkezeti elemzések. Budapest, 1986.
- Beszélt nyelvi tanulmányok. Szerkesztette Kontra Miklós. Budapest, 1988. Linguistica. Series A. Studia et dissertationes, 1.
- Bolla Kálmán: A szöveg fonetikai szerkezetéről. EFF. 2. 1989, 54–95.
- Bolla Kálmán: A hangzásforma akusztikus paraméterei. EFF. 4. 1991, 111–24.
- Bolla Kálmán: A hangminőség mérhetősége mint a kiejtésszabályozás alapja. EFF. 5. 1992, 103–14.
- Bolla Kálmán – Földi Éva: A felolvasás szuprasegmentális szerkezetének fonetikai elemzése. EFF. 5. 1992, 51–63.
- Deme László: Hangtan, Hangsúly, szórend, hanglejtés, szünet. In: MMNyR. Budapest, 1961–2. I, 56–120. II, 458–521.
- Deme László: Szórendi problémák az értekező prózában. In: Műszaki Nyelvőr 1964, 114–35.
- Deme László: A Kazinczy-érmek és a helyes magyar kiejtés kérdése. Nyr. 89: 1965, 188–200.
- Deme László: Mondatszerkezeti sajátosságok gyakorisági vizsgálata. Budapest, 1971.
- Deme László: A bemondói munka mondatfonetikai kérdéseiről. In: Rádbesz. 1973, 71–101.
- Deme László: A szövegalkotás szerkezeti–szerkesztési kérdései. In: Szónokve. 1974, 1975, 116–51.
- Deme László: A rádióban elhangzó beszédéről. MRT TK Tanulmányok VIII, 4. sz. 1976.
- Deme László: A rádióhírek szövegalkata és hangzásformája. RTV Szle. II, 1978, 48–55.
- Deme László: Grammatikai képlet és akusztikai képlet kapcsolatához. MFF. 3. 1979a, 7–13.
- Deme László: A szöveg alaptermészetéről. 1979b. MNyTK. 154. 57–65.
- Deme László: Az ember és a nyelv. In: Anyanyelvi nevelésünk a középiskolában, Budapest, 1980, 9–36.
- Deme László: A szövegség és a szövegegység néhány jellemzője. In: Tanulmányok a mai magyar nyelv szövegtana köréből. Szerk. Rácz Endre – Szathmári István. Budapest, 1983, 31–60.
- Deme László: A szövegmondásról és a szövegről. EFF. 1. 1988, 47–54.
- Deme László: A beszéd, a nyelv, a szöveg. In: Retorika. Budapest, 1988/89.
- Deme László: A szövegszerkezet komplex vizsgálata. EFF. 1. 1989, 25–7.
- Deme László: A hangzásforma hírértéke – a hírérték hangzásformája. EFF. 4. 1991, 75–86.
- Deme László: Szövegalkat és hangzásforma. EFF. 5. 1992, 95–102.
- Dezséry Judit – Terestyéni Tamás: Élő szöveg – stúdiószöveg. ÁNYT. XI, 51–77.

- Elekfi László: Mondatformák és mondathangsúlyok fonetikai vetülete. NyK. 70: 1968, 329–49.
- Elekfi László: Föltámadott a tenger... (Hangsúlyozás és mondanivaló a versben és zenei áttételeiben) Magyar Zene XX, 1979, 391–422.
- Elekfi László: Hanglejtés. NyművKk. Budapest, 1980, I, 747–87.
- Fábricz Károly: A beszélt nyelvi szövegalkotás kérdéseihez. In: BeszNyelviTanulm. 1988, 439–59.
- Fónagy Iván: Füst Milán: Öregség – Dallamfejtés. Budapest, 1974.
- Fónagy Iván – Magdics Klára: Az érzelmek tükröződése a hanglejtésben és a beszédben. NyK. LXV, 1963, 103–36.
- Fónagy Iván – Magdics Klára: A nyomaték jelentésmeghatározó szerepe határozós szerkezetekben. NyK. LXVIII, 1966, 97–117.
- Fónagy Iván – Magdics Klára: A magyar beszéd dallama. Budapest, 1967.
- Hajnal Jenő: A vezércikk mondat- és szövegszerkezeti, -szerkesztési sajátágai. Magiszteriumi dolgozat. Zenta. Kézirat. 1991.
- Kassai Ilona: A magyar beszéd hangsorépítési szabályszerűségei. MFF. 8. 1981, 63–86.
- Keszler Borbála: Kötetlen beszélgetések mondat- és szövegtani vizsgálata. In: Tanulmányok a mai magyar nyelv szövegtana köréből. Szerk.: Rácz Endre és Szathmári István. Budapest, 1983, 164–202.
- Kocsány Piroska: Szövegnyelvészet vagy szövegtípusok nyelvésze? FilKözl. 1989, 26–43.
- [Kontra Miklós:] Beszélt nyelvi tanulmányok. Szerkesztette Kontra Miklós, Budapest, 1988. Linguistica. Series A. Studia et dissertationes, 1.
- Molnár Ildikó – Surányi Ibolya: Művészi versmondások összehasonlító elemzése I–II. In: Tanulmányok a magyar nyelv múltjáról és jelenéről. Benkő- emlékkönyv. Budapest, 1981, 235–54.
- Nagy Ferenc: Bevezetés a magyar nyelv szövegtanába. Egyetemi jegyzet. Budapest, 1981.
- Németh T. Enikő: A hezitációk kiszűrése és a beszédfonetikai elemek szerepe a megnyilatkozás határ-vizsgálatnál. Acta Iuvenum, Sectio Linguistica. T. III, 1987, 24–36.
- Németh T. Enikő: A kötőszók szerepe a megnyilatkozás határ vizsgálatában. Néprajz és Nyelvtudomány 31–2. 1988, 168–81.
- Németh T. Enikő: Az emberi kommunikáció kutatásának néhány alapkérdése. Néprajz és Nyelvtudomány 33. 1989–90, 43–56.
- Péchy Blanka: Beszélni nehéz! Budapest, 1974.
- Petőfi S. János: Szöveg, szövegtan, műelemzés. Budapest, 1990.
- Petőfi S. János: A humán kommunikáció szemiotikai elmélete felé. Szeged, 1991.
- Rádbesz 1973. A rádióbemondó beszéde. Szerk.: Wacha Imre. MRT TK 20. Budapest, 1973.
- Subosits István: A beszédpedagógia alapjai. Budapest, 1982.
- Szabó József: A mondat szerkesztés nyelvészociológiai vizsgálata a nagykönyvi nyelvjárásban. Budapest, 1983.
- Szabó Zoltán: Szövegnyelvészet és stilsztika. Budapest, 1988.
- Szende Tamás: A beszéd folyamat alaptényezői. Budapest, 1976.
- Szónokkve.: Szónokok, előadók kézikönyve. Szerk.: Deme László. Budapest, 1974, 1975.
- Terestyéni Tamás: Az anyanyelvi kommunikációs kultúra néhány jellegzetessége Magyarországon. Budapest, 1987.
- Varga László: A magyar nyelv dallamprozódámai. EFF.2. 1989, 149–215.
- Vass László: Terminológiai szótár (a szemiotikai szövegtan tanulmányozásához) In: Szemiotikai szövegtan. 1. AAPSz SLLA, Szeged, 1990.



- Vértés O. András: Érzelmi világunk és a nyelv történeti változásának kölcsönhatása. MNyTK 179. 1987a.
- Vértés O. András: Bevezetés a magyar hangstilisztikába. NytudÉrt. 124. Budapest, 1987b.
- Wacha Imre: Az elhangzó beszéd szövegfonetikai eszközeinek rendszere és összefüggései. NyK. 75: 1973, 77–103.
- Wacha Imre: Az elhangzó beszéd főbb akusztikus stíluskategóriáiról. ÁNyT. X, 1974. 203–16.
- Wacha Imre: Beszéd: szituáció, szöveg és hangzás együttese a rádióban és a televízióban. In: Nyelvészet és tömegkommunikáció. Szerk.: Grétsy László. Budapest, 1985. I.
- Wacha Imre: Élőnyelvi (spontán) szövegek megnyilatkozásainak (szintaktikai) vizsgálati szempontjaihoz. BeszNyelviTan. 1988, 102–58.
- Wacha Imre: A hangképzés kommunikatív szerepe. a beszéd folyamatban, különös tekintettel a nyelvi rétegződésre és az elhangzó beszéd stíluszintjeire. EFF. 2. 1989, 121–39.
- Wacha Imre: Szempontok a drámai nyelv akusztikumának, akusztikus stílusváltozatainak vizsgálatához. EFF. 3. 1990, 197–208.
- Wacha Imre: Tartalom, szándék és hangzásforma harmóniája. EFF. 4. 1991a, 87–110.
- Wacha Imre: Adalékok a magyar intonáció vizsgálatának történetéhez. In: Tanulmányok a magyar nyelvtudomány történetének témaköréből. Szerk.: Kiss Jenő és Szűts László. Budapest, 1991b, 710–6.
- Wacha Imre: A megnyilatkozáshatár megállapításának gondjaihoz. In: Emlékkönyv Benkő Loránd hetvenedik születésnapjára. Szerk.: Hajdú Mihály és Kiss Jenő. Budapest, 1991c, 650–4.
- Wacha Imre: Szöveg és interpretáció. EFF. 5. 1992a, 65–71.
- Wacha Imre: Nyelvhasználati rétegződés kiejtésben és hangzásban. EFF. 5. 1992b, 81–94.
- Wacha Imre: A retorika vázlata (Kéziratként). A Magyar Rádió Nyelvi Bizottsága, 1992c.
- Wacha Imre: A nyelvi rétegződés kérdései ("Perújrafelvétel" a magyar szakirodalom alapján). In: Normatudat – Nyelvi norma. Szerk.: Kemény Gábor. Linguistica, Series A, Studia et dissertationes, 8. Budapest, 1992d, 49–105.
- Z. Szabó László – Wacha Imre: A Kazinczy-versenyek története. Kazinczy Ferenc Gimnázium, Győr 1978.
- Z. Szabó László – Wacha Imre: A Kazinczy-versenyek huszonöt éve. Kazinczy Ferenc Gimnázium, Győr 1978, 1991.

# GYORSULT-E A MAGYAR BESZÉD TEMPÓJA AZ ELMÚLT 100-120 ÉVBEN?

Kassai Ilona  
MTA Nyelvtudományi Intézet

Hangtani szakirodalmunkban és szakmai eszmecseréken az utóbbi három év-tizedben többen is közölték azt a sejtésüket, hogy a magyar beszéd tempója felgyorsult. A feltevés nem előzmény nélküli, hiszen a beszédíram korszakonkénti gyorsulásának feltételezése számos nyelvtörténeti változásnak megfelelő magyarázó elve lehet (vö. pl. Kubínyi 1958 és 1961).

Tudomásom szerint eddig ketten szolgáltatott közvetett bizonyítékot a sejtésre. Vértés O. András a különböző korokból származó, gyorsírással lejegyzett parlamenti beszédekből számított, a szakirodalomban közölt sebességi értékekre támaszkodva mutatott rá a tendencia érvényesülésére (1987, 38; 1989, 382). Ezeket egészítette ki Szende Tamás egy 1986-ra vonatkozó adattal Baczony Lászlónak, Magyarország "örökös gyorsíróbajnokának" a nyilatkozatára hivatkozva (1987, 46). Az eredeti, szótag/percben közölt szakirodalmi adatokból pótlólag kiszámítottam a tempóindexet a közkeletűbb szótag/mp-ben is, és valamennyi adatot táblázatba rendeztem (1. táblázat):

1. táblázat

Év	Tempóindex	
	szótag/perc	szótag/másodperc
1869	120-140	2,0-2,3
1889	220	3,6
1924	240-250	4,0
1986	250-300	4,2-5,0

Ezek a számadatok látszólag meggyőzően tükrözik a gyorsulás tendenciáját, mégsem maradéktalanul megnyugtatóak. Vértés O. András sem tartja őket annak, sőt maguk a gyorsírók sem (vö. Nosz 1943). A bizonytalansági tényezők a következők. Először is maga az egység, amelyben a tempóindexet a parlamenti gyorsírók megadják, nevezetesen a szótag, csak megközelítőleg képes tükrözni a beszéd iramát. Egyfelől ugyanis a magyar szótagok szerkezete mennyiségileg is, minőségileg is igen változatos, ami miatt pl. a *bánts*d szó szótagja nehezen vehető össze az *öt* szó szótagjával. Másfelől nem derül ki, hogy a szótagot alkotó fonémákból ténylegesen hányat valósított meg a beszélő, a gyorsíró ugyanis normatív helyesírási alakra írja át sajátos jeleit, amelyekkel a hangzó beszédet rögzít.



zíti (vö. Nosz 1963, 97; Szépe 1963, 237). Ebből következik a szótagszámban közölt adatok hitelességének egy további korlátozása. Nevezetesen az, hogy a gyorsírással lejegyzett szöveg egyáltalán nem utal a hangos szakaszokat elválasztó esetleges szünetekre. Ez a tény pedig jelentősen befolyásolja a tempóindexet, hiszen közismert, hogy a beszéd sebességének érzete nemcsak a hangképzés időzítésétől függ, hanem a hangfolyamatot megszakító szünetek számától és időtartamától is, egyéb járulékos tényezőket nem is említve.<sup>1</sup> Előfordul tehát, hogy az egyik országgyűlési képviselő kevesebb szünetet tart, mint a másik, ezért felszólalása során összességében több szótagot produkál, mint képviselőtársa. Annak illusztrálására, hogy ennek a ténynek mekkora jelentősége van, álljanak itt Hegedűs (1957, 227) adatai:

2. táblázat

	AI	h/mp	SzI	Sz <sub>átlag</sub>
hírolvasás	65970	12	24800	708
sportközvetítés	127260	14	15860	323

(AI=artikulációs idő ezredmásodpercben, SzI=szünetidő ezredmásodpercben, h/mp=tempóindex hang/másodpercben kifejezve, Sz<sub>átlag</sub>=átlagos szünet-időtartam ezredmásodpercben).

Az adatokból világosan kiderül, hogy míg a hírolvasásban a szünetek és az artikulációs idő aránya, az előbbi 1-nek véve, 1:2,4, addig a sportközvetítésekben a szünetidőnek nyolcszorosa az artikulációs idő.

A beszédíram felgyorsulásával kapcsolatos másik bizonyítási kísérlet Gósy Mária nevéhez fűződik, aki a tempóészlelés és a beszédmegértés között fennálló összefüggéseket két vizsgálatban igyekezett tisztázni (1988). Az első kísérletben a szerző az artikuláció sebességének és az alaphangfrekvencia modulálásában kifejeződő dallammodulálásnak a tempóészlelésben játszott szerepét vizsgálta egy mesterségesen előállított alapmondat dallamban ötféleképpen, tempóban kétféleképpen módosított változatai alapján. Az eredményekből egyikként azt a következtetést vonja le, hogy "az észlelés szintjén legalábbis bizonyítható beszédünk felgyorsulása" (103). Következtetését arra alapozza, hogy a kísérletben részt vevő hallgatók azt a mondatváltozatot minősítették normálnak, amelyiket az alapváltozat tempójához képest 1,2-szeres gyorsításban hallottak. Ha azonban tüzetesen megvizsgáljuk a táblázatban közölt számadatokat (i.m. 102), azt látjuk, hogy az alapmondat tempóindexe 9,2 hang/mp, ami Fónagy Iván és Magdics Klára

<sup>1</sup> A tempószámítás két lehetséges módja a beszédhangok képzésére fordított idő, másfelől a tulajdonképpeni közlésíram számítása, amely utóbbi a közléstartamba beleérti a szünetet, hezitációt, tehát minden, nem a beszédképzésre fordított időjelenséget.

mérései szerint 1960-ban a versmondást jellemezte (vö. 3. táblázat). Az 1,2-szeresére gyorsított változatban szereplő tempóindex 11 hang/mp, ez lényegében megegyezik a Fónagy – Magdics-féle köznyelvi átlaggal, a 11,35 hang/mp-cel (1960, 457). Nem véletlen tehát, hogy éppen ezt a mondatváltozatot minősítették a hallgatók normál tempójúnak. Mivel pedig az 1,6-szeresére gyorsított harmadik változat gyorsnak minősült az észlelés számára, leszögezhetjük, hogy a 15 hang/mp-es tempóindex, amely ezt a mondatot jellemezte, magasnak tekinthető a mai nyelvhasználatban. Az idézett kísérlet tehát, ha helyesen értelmezzük az adatokat, voltaképpen nem visz közelebb a címben feltett kérdés megválaszolásához.

3. táblázat

Beszédműfaj	A beszédsebesség középértéke beszédhang/mp
Vers	9,4
Mese	9,57
Prózaolvasás	10,73
Gyerekbeszéd	11,68
Beszélgetés	12,89
Sportközvetítés	13,83
Középérték	11,35

A tempógyorsulásra vonatkozó sejtés közvetett úton való igazolására ritka, megbízható támpontot kínál Gombocz Zoltán 1909-ből származó, egyre növekvő szótagszámú szósorozattal végzett, híressé vált kísérlete, a *tát, tátog, tátogat, tátogatók* és *tátogatóknak*, amelyet Tarnóczy Tamás 1965-ben megismételt.<sup>2</sup> A két kísérlet újbóli felhasználása aktuális kérdésünk szempontjából azért ideális, mert (1) a szünetek tempóészlelést befolyásoló hatása kiküszöbölődik, (2) a nyelvi anyag mindkét esetben azonos, (3) a bemondó mindkét esetben hasonló életkorú, iskolai végzettségű és társadalmi helyzetű férfi, (4) a hangfelvétel laboratóriumi körülményei kizárták a redukált artikuláció (csökkentett mennyiségű artikulációs mozgás) lehetőségét, végül (5) a két felvétel és értékelés között mintegy 60 év, tehát két emberöltőnyi idő telt el. Legalább részben teljesülnek tehát azok a feltételek, amelyek mellett Szende bizonyíthatónak tartja a tempógyorsulást: "Valamelyest is hiteles megítélés csak úgy lehetséges, ha tematikailag nagyjából megegyező, nagyjából azonos közlési helyzetben és nagyobb történeti távlatban adott, statisztikailag érvényes eredményeket ígérő, tehát nagyobb közlésmennyi-

<sup>2</sup> Tarnóczy az automatikus, gépi beszédfelismeréssel kapcsolatban igyekezett tisztázni az időtartamok alapján elvégezhető hangzófelismerés elvi alapjait, és ehhez használta fel a kísérletet. Eredményeit angolul 1965-ben, magyarul 1966-ban és részben 1974-ben tette közzé.



ségre épülő számításokat, esetleg megbízható becsléseket látunk" (1987, 19). A két kísérlet egymásra vetített táblázatának adatai ezért fontos következtetéseket tesznek lehetővé (vö. 4. táblázat).

4. táblázat

Szavak	A magánhangzók időtartama (ms)									
	a:		o		ɔ		o:		ɔ	
	1909	1965	1909	1965	1909	1965	1909	1965	1909	1965
tát	272	210								
tátog	242	180	118	155						
tátogat	209	140	94	95	109	115				
tátogatók	190	120	79	85	87	105	221	130		
tátogatóknak	182	110	79	80	86	90	172	110	–	140

Ha a számoszlopokat szemléljük, megállapíthatjuk, hogy minél több szótagból áll a hangsor, annál rövidebb az egyedi hangok időtartama. Ha pedig a sorokat nézzük, az tűnik ki, hogy minél távolabb van egy hang a hangsor elejétől, annál hosszabb időtartammal valósul meg. Elég, ha megnézzük a 3. és az 5. szótagban szereplő [ɔ] hangok időtartamértékeit: az 5. szótagbeli [ɔ] 55%-kal nagyobb időtartamú, mint 3. szótagban szereplő párja. De ugyanez a hang 27%-kal még az 1. szótagban szereplő, fonológiai hosszú /a:/ időtartamát is meghaladja. Mármint a két szerző adatainak a közvetlen összevetésével kimutatható, hogy a századelő óta mennyire közeledtek egymáshoz a rövid és a hosszú magánhangzók időtartamértékei. A Gombocznál megfigyelhető, o/o: párra vonatkozó 1:2,79, illetőleg 1:2,17-os időtartamaránnyal Tarnóczy adatai szerint 1:1,52 és 1:1,37 áll szemben. Ugyancsak figyelemre méltó, hogy a 2. és a 3. szótagban Tarnóczy abszolút időtartamértékei meghaladják a Gombocz által közölt értékeket.<sup>3</sup> Ez úgy interpretálható, hogy az ejtési időtartamok nemcsak a hosszú magánhangzók erőteljes rövidülése révén kerültek közelebb egymáshoz, hanem a rövid magánhangzók nyúlása révén is.<sup>4</sup> Nyilvánvaló azonban, hogy a hosszúak rövidülése az erőbben ható tendencia. Az adatok további értelmezéséhez Mag-

<sup>3</sup> Ennek pedig annál is nagyobb a heurisztikus értéke, mivel Gombocz a hangelhatárolásnál másként járt el, mint Tarnóczy: a zárhangok zárképzési szakaszát a megelőző magánhangzóhoz, a zár-felpattanási szakaszát pedig a következő magánhangzóhoz számította (Meyer – Gombocz 1909, 5.). Vagyis a magánhangzókra általa közölt adatok eleve több információt tartalmaznak, mint Tarnóczy megfelelő adatai.

<sup>4</sup> Ugyanezt a tendenciát állapítottam meg az időtartam alakulására vonatkozó átfogó vizsgálatomban: az időtartamot befolyásoló tényezők által kifejtett nyújtó hatás a rövid beszédhangokat érintette nagyobb mértékben, a rövidítő hatás pedig a hosszú beszédhangokat (Kassai 1979, különösen 29–36).

dics Klára tapasztalataira utalok, akinek a beszédtempó időtartammódosító hatásával kapcsolatos vizsgálatából (1969) legfontosabbként az derült ki, hogy minél hosszabb az egyes magánhangzók időtartama nyugodt beszédtempó mellett, annál fokozottabb rövidülés mutatkozik a gyors tempóra való áttéréskor. Az összefüggés fordítva is áll. Mármint, mivel a mi esetünkben a két táblázat tanúsága szerint a hosszabb időtartamú hangok erőteljesebben rövidültek a két vizsgálat között eltelt hat évtized folyamán, a felszínre hozott tendencia csakis azzal magyarázható, hogy a beszédtempó a század elejéhez képest felgyorsult. Az érvelés más vizsgálati területen is megerősítést nyer. A franciaországi telekommunikációs kutatás fellelőjében, Lannionban a francia mesterséges beszéd létrehozói arról számoltak be, hogy a mesterségesen előállított hosszabb szavakban, szövegekben a hangsor vége felé fokozatos tempógyorsulás észlelhető. A jelenség magyarázata jó valószínűséggel a következő. A CNET<sup>5</sup> mesterséges beszélő rendszere előregyártott hangszeletekből, úgynevezett difonokból épül fel, amelyeknek az alapparaméterei, köztük az időtartam, egy külön program segítségével, az előre meghatározott minimum- és maximumérték között változtathatók attól függően, hogy a létrehozandó mesterséges hangsor mely pontján használják fel ezeket az építőelemeket (vö. Stella 1985, 438–450; Bartkova – Sorin 1987, Bartkova 1990). Ha ez így van, a természetes nyelvben a hangsorterjedelem növekedésével együttjáró időzítési jelenségek csak korlátozottan érvényesülnek, ezáltal megtévesztik az észleletet: a gyorsulás érzetét keltik. Ebből viszont az következik, hogy a beszédtempó gyorsulása a szükségesnél rövidebb időtartammal függ össze.

Érthető ezek után a kíváncsiság, hogy milyen kép kerekedik ki a Gombocz-féle kísérletből közel egy újabb emberöltő elteltével. 1992-ben megismételtem tehát az eljárást a minden egyéb feltétel azonosságának a követelményét betartva. Az adatokat hozzárendeltem az előző két vizsgálat adataihoz, így jött létre az 5. táblázatban szemlélhető összesítés. Az 1992-es adatok megerősítik azokat az összefüggéseket, amelyeket az előző két vizsgálattal kapcsolatban feltártam, és a közvetlen szemlélet számára is jól kivehetően dokumentálják az időtartamok további rövidülését. Emellett ha szótagról szótagra kiszámítjuk a szomszédos magánhangzók időtartamának egymáshoz képesti arányát (6. táblázat), számszerűen is bizonyítjuk a rövid és a hosszú magánhangzók időtartamának egymáshoz való közeledését, azaz az időtartamértékeknek egy központi érték felé való tartását. Az értékeknek ez a centralizálódása csökkenti az időzítésbeli változatosságot, a nyomában kialakuló monotonia pedig hozzájárul a tempógyorsulás érzetéhez. A szubjektív benyomás azon alapszik, hogy mivel az észlelési mechanizmusnak egyneműbb információt kell adott időpillanatban feldolgoznia, a feldolgozás menete felgyorsul.

<sup>5</sup> Centre National d'Etudes de Télécommunication.



5. táblázat

Szavak	A magánhangzók időtartama (ms)														
	a:			o			ɔ			o:			ɔ		
	1909	1965	1992	1909	1965	1992	1909	1965	1992	1909	1965	1992	1909	1965	1992
tát	272	210	227												
tátog	242	180	176	118	155	110									
tátogat	209	140	141	94	95	78	109	115	102						
tátogatók	190	120	117	79	85	58	87	105	94	221	130	117			
tátogatóknak	182	110	117	79	80	55	86	90	78	172	110	102	–	140	94

A hangsorterjedelem és az időtartam összefüggése 1909-ben, 1965-ben és 1992-ben.

6. táblázat

Szavak	A magánhangzók időtartama (ms)											
	a:/o			o/ɔ			ɔ/o			o:/ɔ		
	1909	1965	1992	1909	1965	1992	1909	1965	1992	1909	1965	1992
tátog	0,49	0,86	0,63									
tátogat	0,45	0,68	0,55	1,16	1,21	1,31						
tátogatók	0,42	0,71	0,50	1,10	1,24	1,62	2,54	1,24	1,24			
tátogatóknak	0,43	0,73	0,47	1,09	1,13	1,42	2,00	1,22	1,31	–	1,27	0,92

A szomszédos szótagok magánhangzóinak időtartamaránya %-ban

A három kísérlet adatainak az elemzése alapján kimondható, hogy a magyar beszéd tempója valóban felgyorsult, legalábbis a század elejéhez képest. A gyorsulás élménye alapvetően két tényezőhöz köthető. Első fokon az abszolút időtartamértékek csökkenése értelmeződik tempógyorsulásként, áttételesen pedig a belső arányok átrendeződése nyomán kialakuló időzítési monotonía kelti a növekvő sebesség érzetét. Arra a kérdésre, hogy mi az átlagos köznapi beszéd tempóindexe ma, az egyik lehetséges választ Gósy Mária hivatkozott vizsgálatának második kísérletéből kaphatjuk meg, amely a beszédtempónak a beszéd megértésére gyakorolt hatását kutatta. Ebben a természetes ejtésű beszédanyagban ugyanis a normál tempó képviselésében 13,98, kerekítve 14 hang/mp jelenik meg (vö. 7. táblázat), ami lényegében azonos a Fónagy – Magdics (i.m. 457) és a Hegedűs (i.m. 227) által jelzett, jó 30 évvel ezelőtti sportközvetítések tempóindexével.

7. táblázat

Észlelt tempó	Tempóindex		Tempóváltozás (%) $\frac{\Delta T}{T}$
	h/mp	szótag/mp	
normál	13,98	5,97	—
gyors	23,68	10,16	59
nagyon gyors	29,52	12,69	47
nagyon lassú	7,2	3	194

### Irodalom

- Az élőszó nyomán. Szerk. Nosz Gyula. Budapest 1943.
- Bartkova, K. – Sorin, Ch.: A model of segmental duration for speech synthesis in French. *Speech Communication* 6. 1987, 245–260.
- Bartkova, K.: Règles du modèle prédictif de la durée segmentale. Document de travail 1095/LAA/TSS/RCP. 1990. [Kézirat.]
- Fónagy Iván – Magdics Klára: Beszédsebesség, szólam, ritmusérzék. *MNy* 56. 1960, 450–458.
- Gombocz Zoltán: A magyar hangok időtartamáról. *Nyelvtudomány* 2. 1909, 93–100.
- Gósy Mária: Tempóérzékelés és beszédmegértés. *Műhelymunkák* V. 1988, 87–122.
- Hegedűs Lajos: Beszédtempó-elemzések. *Nyr* 81. 1957, 223–227.
- Kassai Ilona: Időtartam és kvantitás a magyar nyelvben. *NytudÉrt* 102. 1979, Budapest.
- Kubinyi László: A beszédíram magyar nyelvtörténeti jelentőségéről. *MNy* 57. 1961, 337–340.
- Kubinyi László: Magyar nyelvtörténeti változások vélhető összefüggéséről. *MNy* 54. 1958, 213–32.
- Magdics Klára: A magyar beszédhangok időtartama nyugodt és gyors beszédben. *NytudÉrt* 67. 1969, 45–63.
- Meyer, E. A. – Gombocz, Z.: *Zur Phonetik der ungarischen Sprache*. Uppsala 1909. (Különlenyomat a *Le Monde Orientale* 1907-08. évi számából.)
- Nosz Gyula: A gyorsíró és az élőszó. *Nyr* 87. 1963, 96–104.



- Stella, M.: Speech synthesis. in: Computer Speech Processing. Szerk. Fallside, F. – Woods, W. A. Cambridge 1985, 421–460.
- Szende Tamás: Megérthetjük-e egymást? Budapest 1987.
- Szépe György: Nyelvészeti jegyzetek a beszédgyorsírásról. Nyr 87. 1963, 231–237.
- Tarnóczy Tamás: A magánhangzók akusztikai vizsgálatának problémái. ÁltNyT 10. 1974, 181–196.
- Tarnóczy Tamás: Can the problem of automatic speech recognition be solved by analysis alone? Rapport du 5<sup>e</sup> Congrès Int. d'Acoustique II. Conférences générales. Liège 1965, 371–387.
- Tarnóczy Tamás: Megoldható-e az önműködő beszédfölismerés kizárólag elemzés alapján? Fizikai Szemle 16. 1966, 133–139.
- Vértés O. András: Érzelmi világunk és a nyelv történeti változásának kölcsönhatása. MNy 82. 1986, 11–29.
- Vértés, O. A.: Gedanken über die geschichtliche Veränderung des Sprechtempos. Proceedings of the Speech Research '89 International Conference. MFF/HPP 21. 1989, 381–384.
- Vértés O. András: Időtényezők stílusbeli szerepe. Műhelymunkák V. 1988, 123–147.

# SZEMÉLYI SZÁMÍTÓGÉP ÉS FONETIKAI-FONOLÓGIAI KUTATÁSOK

Papp Ferenc

MTA Nyelvtudományi Intézete, Budapest

PC-ket nevezett kutatásokra több módon lehet alkalmazni. Itt két irányra szeretnék rámutatni: azoknak a szavaknak a teljes listázása, melyek ún. archifonémákat meg azonos magánhangzó fonémakettősöket tartalmaznak (1.); a hosszú valamint az álhosszú többjegyű betűket tartalmazó címszavak listájának hasonló összeállítása (2.). Példáim mindkét esetben az ÉrtSz.-re, tehát egy lexikográfiai korpuszra fognak korlátozódni. Természetesen semmi akadálya sincs annak, hogy ugyanezt a munkát szövegeken, így az MTA Nyelvtudományi Intézete Nagyszótári Osztályán készülő hatalmas szöveg-thészauruszon elvégezzük. Esetenként utalni fogunk erre a lehetőségre s annak sajátosságaira. Megjegyzem még, hogy a két irányt egyesíti azok hasonló felszíni – tehát pl. számítógépes – megoldása: adott karaktertöbbsort keresünk a file-ban: ez természetesen rutinfeladat (miután KG a technikai-programozási kulcsot megtalálta hozzá).

**1.1. gyöngy, pinty.** Tekintsük meg ezt a két, itt fettel szedetett tőszót, az 1. sz. függelékben bővebben adott válogatást. Bennük természetesen nem az a probléma, amit a HSzab.<sup>11</sup> említ, hogy ti. az egyszerűsítés elve érvényesül leírásukban, bár bizonyos értelemben ez a primitív és félrevezető meghatározás is utal valamire. Nevezetesen: ilyen esetekben kétségtelenül [ɲ] hangot ejtünk: gyö[ɲɟ], pi[ɲc]. Ezt egyszerűsítve "n" betűvel írjuk – ámde jól tudjuk, hogy betűkkel nem hangokat, hanem fonémákat jelölünk. Milyen fonéma van ezeken a helyeken, /n/ vagy /ɲ/, ennek megfelelően mit kell itt írunk, "n" vagy "ny" betűt? A természetes ejtéshez (hangokhoz) közelebb álló fonológiai irányzatok számára olyan fonéma van, amilyen hangot ejtünk – esetünkben tehát / /, eszerint "ny" betűt is kellene írunk. Más irányzatok megkövetelik, hogy a kérdéses hangot szabad helyzetében vizsgáljuk. Amilyen hangot ott hallunk, olyan fonémával van dolgunk: *vasgolyó* – /ʃ/ fonéma, mert szabad helyzetben [vɔʃ], s-sel; *vagongyár* /n/ fonéma, mert [vɔgon] [n]-nel stb., függetlenül az aktuális [ʒ], [ɲ] stb. ejtéstől.

Könnyű volt a helyzetünk a fenti esetekben s a hozzájuk hasonló összetett szavakban: az összetételt, képzős alakulatot részeire bonthatjuk s így "kiugrathatjuk" az előtag (a tő) végső mássalhangzójának szabad karakterét. De mit tehetünk a fejezet címébe foglalt s ahhoz hasonló tőszavakban? Hiszen ezek azért tőszavak, hogy felbonthatatlanok, nincs mit kiugratni: a [ɲ] ezekben bizony mindig



csak [ɲ], nem tudhatjuk, hogy a /n/ fonéma realizációi-e, avagy a ɲ-éi. (Ezért kvázi mindegy, hogy mit írunk, ezért lehet éppen itt "egyszerűsíteni" és "n"-et írni és nem az ezer és ezer más helyen, ahol persze kötelezően "ny"-et vetünk papírra-táblára. Ez egyeseknek borzasztó, de így van: a magyar helyesírás elég következetesen fonématis. Csak "elég következetesen" és nem teljesen – egy természetes úton kialakult rendszertől ne várjunk természetellenes, teljes következetességet.) (Ezt a kérdést l. még Hell kandidátusi értekezésében: kézirat [1983], 127, 131.)

Van ennek a, mint látjuk, éterien elméleti kérdésnek egy nagyon gyakorlati vonatkozása. Nevezetesen: a helyesírás, kivált a magyar gyermekeké. (Máskor a magyar gyermekekkel egy sorban fogjuk emlegetni a magyarul tanuló külföldit. Itt ezt nem tehetjük, legalábbis ha könyvből mesterségesen tanuló külföldit tartunk szem előtt: az ilyen idegenajkú előbb tanulja meg a betűt, s végig az marad az uralkodó számára.) Milyen hibákat vét a magyar kisiskolás a *gyöngy*, *pinty*-féle szavakban? Amikor a társadalmilag helyes fonématis írásmódot beleverjük az iskolában – mennyire verjük ki belőle a jó fonetikus hallást? Vagy inkább az a helyzet, hogy mire ezek a ritkább szavak gyermekünk elé kerülnek, az már sziklaszilárdan ismeri a társadalmilag elvárt norma elveit, s ezért – a magyarban – megszűnt becsületesen hallani? A gimnáziumra alkalmasint az a helyzet, hogy e szavak nem okoznak helyesírási problémát?

Éppen a magyar helyesírás fonématis volta miatt még néhány hasonló elem megbújhat a *ndj*, *ntj* stringek képe alatt is. Ezek nem lévén túl számosak, itt fel is sorolom őket. *ndj*: *mindjárt*, *mindjobban*, *rendjel*; *ntj*: *elrikkantja magát*, *szentjánosáldás*, *-bogár*, *-kenyér*, *szintjel*, *-zés*, *-ő*. De ezek az ÉrtSz.-címszavak talán csak egy kivétellel (*mindjárt*) egészen nyilvánvalóan, szigorúan szinkron szempontból is összetettek vagy kvázi szövegből vett nagyobb darabok. Az előbbiekre nyilván más szabályok érvényesek (vö. legutóbb pl. Deme 1992, 30), az utóbbiak szinte maguktól felhívják a figyelmet arra, hogy szövegtörzsekben nem csupán más statisztikus törvényszerűségek fognak uralkodni (ez természetes), de más grammatikai-morfológiai helyzetek is fognak szerepelni (*mondja*, *pontja*; [minden] *szentje* stb.), emiatt szövegekben várhatóan gyakoribbak is. Akár így, akár úgy alakul azonban, a szövegek prezentálta ilyen stringek feltehetően nem lesznek archifonématisak – így nevezhetjük a *gyöngy*, *pinty*-féle feloldhatatlan tőszói alakulatokat –, alkalmasint mind átlátszó morfológiai okoknál fogva, a fonématis elveknek megfelelően írónak így vagy úgy.

**1.2.** Korpuszainkban, lexikográfiaiakban éppenúgy, mint élő szövegen, előfordul, hogy két azonos magánhangzó-fonéma kerül közvetlenül egymás mellé (*aláás*, *beebédel* stb. – l. 2. sz. függelék). Hogy itt két fonémáról van szó, azt világosan mutatja mind a 15 milliónk nyelvérzéke még akkor is, ha gyors, hanganyag ejtéssel egészen egybe is rántjuk a két (azonos) magánhangzó-elemet<sup>1</sup>. Az

osztatlan hangpatak fonémákra bontása éppen a szociális faktor dolga, az nyelvenként-hagyományonként esetleg más és más lehet, ezért hivatkoztunk itt a 15 millió magyarra, hazánk határain belül és kívül. Ámde ha megnézzük Olasz Gáborék műszerrajzolta görbéit, ijedten látjuk, hogy e helyen nincs szakadás, amint várnók – valami tehát társadalmilag (fonológiaiilag) van is, meg – fizikailag – nincs. (Vö. Olasz 1989, 63.) Vagyis a "Fonnyadt babarcú öregasszony" és a "Gyönyörű, babarcú fiatalasszony" homofónok, de nem homográfok – csak ennyi lenne köztük a különbség? Vajon találunk-e még ehhez hasonló minimális párokat szótári vagy szöveganyagban?

Itt közlöm az ÉrtSz. e kettősöket tartalmazó címszavainak teljes táblázatos áttekintését; a hivatkozott függelékben hozom a népesebb teljes listáikat.

fon. kettős	Σ	100=183	lista
aa	47	25,68	babaarc, bábaasszony, bérmaanya, bérmaapa, faanyag, faarc, hasaalja... <i>teljes listát l. a 2.sz. függelékben</i>
áá	5	32,73	aláás, hozzááll, hozzáállás, rááldoz, rááll
ee	74	40,44	bcebédél, beecsetel..., beleegyezés, bűnbeesés, csereeszköz, csirkeeszű, lcemel, ... zeneegylet <i>teljes listát l. i. h.</i>
éé			0
ii	35	19,13	antiimperialista, férűing stb. <i>teljes listát l. i. h.</i>
íí			0
oo	10	5,46	kooperáció, kooperál, kooptál, koordináció, koordinál, koordináta, koordináta-rendszer, mikroorganizmus, zoológia, zoológus
óó	4	2,19	állóóra, blokkolóóra, disznóól, fogadóóra
öö			0
őő	6	3,28	előőrs, erdőőr, mezőőr, úttörőőrs, védőőrizet, védőőrség
uu	2	1,09	individuum, vákuum
úú			0
üü			0
űű			0
	183	100,00	

**2. asszony, csecsszopó:** hosszú és álhosszú magyar mshzó fonémákat tartalmazó címszók kigyűjtése. Az ÉrtSz. keretei között ezt megtettem a szegedi előadásomban, itt onnan kiemelem a fő táblázatot:



	/XXX/	/X-XY/	$\Sigma$ /string	/X-XY/ példa/teljes felsorolás
ccs	88,89 40 2,44	11,11 5 2,75	<b>45</b> 2,47	arccsont, gerinccsatorna, -cserép, harccsoport, tánc-
ggy	41,84 41 2,50	58,16 57 31,32	<b>98</b> 5,38	anyaggyűjtés, bélyeggyűjte- mény, ... foggyökér, ideggyen- ge, jéggyár, joggyakorlat
lly	100,00 31 1,89	0,00 0 0,00	<b>31</b> 1,70	<b>0</b>
nny	93,57 160 9,77	6,43 11 6,04	<b>171</b> 9,40	agyonnyom, ellennyugta, ... is- tennyila, ... különnyomat, negy- vennyolcas, színnyomás ...
ssz	92,18 1250 76,31	7,82 106 58,24	<b>1356</b> 74,51	adásszünet, ... gyorszámoló ... [79], csecsszopó [13], ésszerű [5], törzsszám [5], össz-szövet- ségi [4]
tty	100,00 116 7,08	0,00 0 0,00	<b>118</b> 6,37	<b>0</b>
zzs	0,00 0 0,00	100,00 3 0,65	<b>3</b> 0,16	halászsákmány, horgászszi- neg, vadászsákmány
$\Sigma/..$	<b>1638</b> 90,00	<b>182</b> 10,00	<b>1820</b> 100,00	

És külön a (d)dz(s) stringgel kapcsolatos adatok:

	/dz/	/d-z/	$\Sigma$ /string	teljes felsorolás
dz	95,28 121 79,08	4,72 6 100,00	<b>127</b> 79,87	beszédzavar, csendzavarás, földzárlat, rendzavarás, smaragdzóld, vadzab
ddz	0	0	<b>0</b> 0,00	<b>0</b>
dzs	100,00 32 20,92	0,00 0 0,00	<b>32</b> 20,13	[tehát <b>*VADZSEB</b> nincs!]
ddzs	0	0	<b>0</b> 0,00	<b>0</b>
$\Sigma/..$	<b>153</b> 96,23	<b>6</b> 3,77	<b>159</b> 100,00	

Ezek a példák elméletileg is, de talán kivált gyakorlatilag érdekesek. Hogyan is olvasunk, amikor magyarul olvasunk? Ha mi, jól olvasó magyar felnőttek, már el is felejtettük ezt (= az olvasás folyamata automatikussá vált bennünk) – ám hogyan olvasnak gyermekeink/unokáink, hogyan olvasnak magyar szövegeket a magyarul tanuló külföldiek, hiszen ebben a két csoportban még többé-kevésbé nem automatizáltak ezek a folyamatok. Én itt természetesen nem kívánok erre a kérdésre válaszolni. Csak, előadásom címének megfelelően, nyersanyagot kívántam szolgáltatni efféle kísérletek elvégzéséhez, felhívni a figyelmet arra, még mennyi efféle nyersanyag áll a rendelkezésünkre géppel könnyen összegyűjthető módon mind a szótárakban, mind a szövegekben. – És még egy, részint ezzel összefüggő, gyakorlati kérdés – az elválasztásé.

A magyar szóelválasztás – egyébként kár büszkélkednünk: ilyesfajta még a német, a francia, az orosz stb. (csak nem az angol!) – olyan transzparens, hogy szinte magától kínálta magát az automatikus elvégzésre, az algoritmizálásra. Évtizedekkel ezelőtt éppen ezért ezen tanítottunk magyar szakos hallgatókat programozni (már elfelejtettem, de – *horribile dictu* – lehet, hogy gépi nyelven). Feltehetően ugyanezért nem sokkal utánunk a berlini akadémia nyelvtudományi intézetében ugyanezt a magyar nyelvi feladatot megoldották. Ez még hagyján – a magyar nyelv/helyesírás mondott sajátosságán kívül itt közrejátszhatott esetleg az is, hogy nekünk, magyaroknak már akkor is megalapozottan jó hírünk volt a táboron belül, a kétfajta németek között pláne... De azután alkalmasint ugyanezért az angolok csak úgy adtak el nekünk egy nagy nyomdai berendezést, ha vele együtt megvesszük a magyar elválasztás őáltaluk összeállított szabályait. És a furcsa eredmény – illetőleg "eredmény" –: egyéb külföldi partnereink (legutóbb mintha egy olasz cégről hallottam volna) csak így hajlandók szállítani "író" – nyomdai, feliratozó stb. – berendezéseiket. Ugyanakkor valahol elveszhetett útközben az alapötlet, a magyar szóelválasztás könnyű automatizálhatósága. Állítólag ennek a körülménynek (a magyar elválasztás szabályai nem-ismeretének) köszönhetjük a minden oldalról, így a magyar sajtóból is hirtelen ránczúduló rettenetes elválasztások tömegét. (Ti. sajtónk is import nyomdagépeken dolgozik.) És mi lenne a helyzet, ha a kérdés m a g v á t évtizedekkel ezelőtt nem oldottuk volna már meg? (Éppen csak a magvát. Az összetételek helyét a memóriában csupán jeleztük, tehát maradt *\*kö-rút* és *\*ren-dőr*, mindig csak *e-lég* és *fe-lül*.)



**1. sz. függelék:** a gyöngy-, pinty-féle szavak (bekezdve a tőszavak):

(Az egyes oszlopok értelme: a) hány tő; b) három oszlop együtt: 001 ige, 002 fn, 003 mn, két szám – kettős szófajúság; c) képző: 0 nincs képző, 1 képző a szó végén, 2 képző a szó belsejében, 3 itt is, ott is, 9 különös képző, rag mint képző stb.)<sup>2</sup>

**ngy (135 rekord, 15 tőszó)**

angyal	10020	gyöngyhímzés	20021
angyalarcú	20031	gyöngyírás	20021
angyalbögyörő	20020	gyöngykagyló	20020
angyalbőr	20020	gyöngykór	20020
angyalcsináló	20321	gyöngyköles	20020
angyalfia	20029	gyöngyös	10321
angyalhaj	20020	gyöngyözik	10011
angyali	10031	gyöngysor	20020
angyalka	10021	gyöngyszem	20020
angyalos	10031	gyöngyszín	20230
arkangyal	20020	gyöngyszínű	20031
ángy	10020	gyöngytyúk	20020
ángyi	10021	gyöngyvászón	20020
árpagyöngy	20020	gyöngyvirág	20020
babakelengye	20020	hangya	10020
begöngyöl	60011	hangyaboly	20020
békeangyal	20020	hangyamaszás	20021
bongyor	10030	hangyasav	20020
csecsemőkelengye	20020	hangyaszorgalom	20021
daragyöngy	20020	hangyatozás	20021
drágagyöngy	20020	hangyász	10021
elrongyol	60011	igazgyöngy	20020
elrongyolódik	60011	ingen	16329
fagyöngy	20020	ingenyes	10329
felgöngyöl	60011	ingenélő	20329
felgöngyölít	60011	ingenhely	20029
félingyenes	20329	ingenjegy	20029
géprongy	20020	ingyért	10069
göngyöl	10011	kalangya	10020
göngyöleg	10021	kalangyál	10011
göngyölít	10011	kapcarongy	20020
göngysúly	20020	kelengye	10020
göröngy	10020	kengyel	10020
göröngyös	10031	kengyelfutó	20021
gyöngy	10230	kengyelszár	20020
gyöngyárpa	20020	kengyelszj	20020
gyöngybagoly	20020	kengyelvas	20020
gyöngybetű	20020	kigöngyöl	60011
gyöngyélet	20021	kigöngyölít	60011
gyöngyfonál	20021	kigöngyölődik	60N11
gyöngyfüggöny	20021	langy	10030
gyöngyfüzér	20021	langymeleg	20030
gyöngyhalász	20021	langyos	10031
gyöngyharman	20020	langyosít	10011
gyöngyház	20020	langyosodik	10M11

langyosság	10021	rongybaba	20020
lengyel	10230	rongylabda	20020
lengyeles	10031	rongyol	10011
lengyelke	10021	rongyolódik	10011
lengyolcs	20020	rongyos	10321
lerongyolódik	60011	rongyoskodik	10011
lerongyolt	60031	rongyosleves	20023
öngyilkos	20329	rongyosodik	10N11
öngyilkosjelölt	30329	rongypapír	20020
öngyilkosság	20029	rongszedő	20321
öngyulladás	20029	rongszönyeg	20021
öngyújtó	20029	spungya	10020
összegöngyöl	60011	szappangyökér	20020
őrangyal	20020	toprongyos	10031
papírrongy	20020	törlőrongy	20022
pongyola	10230	üveggyöngy	20020
pongyolaság	10021	vagongyár	20020
porrongy	20020	varangy	10021
rézangyal	20020	vándorkelengye	20020
ringy-rongy	20230	védangyal	20020
ringyó	10020	védőangyal	20022
rongy	10230	vízgyöngy	20020
rongyász	10021		

### nty (43 rekord, 8 tőszó)

arany-sarkantyús	20031	kulipintyó	10020
békaporonty	20020	lappantyú	10020
billentyű	10021	leffentyű	10021
billentyűs	10031	megkontyol	60011
billentyűzet	10021	megsarkantyúz	60011
csappantyú	10021	pattantyú	10021
díszpinty	20020	pattantyús	10231
elszontyolódik	60011	pinty	10020
felkontyol	60011	pintyőke	10021
felkontyoz	60011	ponty	10020
fogantyú	10021	poronty	10020
herkentyű	10020	röppentyű	10021
hópinty	20020	sarkantyú	10021
kallantyú	10021	sarkantyúka	10021
kallantyús	10031	sarkantyús	10031
karmantyú	10021	sarkantyúvirág	20022
kattantyú	10021	sarkantyúz	10011
konty	10020	szellentyű	10021
kontyalávaló	30029	szívbillentyű	20021
kontyos	10321	tükörponty	20020
kontyvirág	20020	zongorabillentyű	20021
kösöntyű	10020		



**2. sz. függelék:** Az ÉrtSz. 'aa'-t, 'ee'-t és 'ii'-t tartalmazó címszavainak teljes listája. (A két rövidebb csak a címszavakat tartalmazza.)

### aa (47 rekord)

babaarc	munkaasztal
bábaasszony	odaad
bérmaanya	odaadás
bérmaapa	odaadó
faanyag	odaajándékoz
faarc	óraadás
hasaalja	óraadó
hazaad	pályaavató
hálaadás	példaadás
hálaadó	propagandaanyag
hátraarc	puskaagy
katonaanyag	rakétaautó
katonaasszony	rostaalj
kártyaadósság	ruhaakasztó
kártyaasztal	ruhaanyag
kornaasszony	szobaasszony
konyhaajtó	tapétaajtó
koponyaalap	utcaablak
libaaprólék	utcaajtó
macskaasztal	visszaad
mostohaanya	visszaakaszt
mostohaapa	visszaalakul
munkaadó	vitaanyag
munkaalkalom	

### ee (74 rekord)

beebédél	60011	bűnbeesés	20029
beecsetel	60011	csereeszköz	20020
beemel	60010	csirkeeszű	20031
beemléz	60011	egybeesik	60019
beenged	60011	félreesik	60019
beerdősít	60011	félreeső	60039
beereszt	60011	kétségbeejt	20019
beeresztő	60321	kétségbeejtő	20039
beerősít	60011	kétségbeesett	20329
beesett	60031	kétségbeesés	20029
beesik	60010	kétségbeesik	20019
beesteledik	60M11	közbeesik	60019
beeszik	60010	közbeeső	60039
beevéz	60010	leegyszerűsít	60011
beleegyezés	60029	leegyszerűsödik	60011
beleegyezik	60019	leejt	60010
beleelegyedik	60019	leellik	60010
beleelegyül	60019	leelőlegez	60019
beleereszt	60019	leemel	60010
belesik	60019	leendő	10031
beleeszik	60019	leendőbeli	10321

leenged	60011	összeerősít	60011
leereszkedés	60021	összeesés	60021
leereszkedik	60011	összeesik	60010
leereszkedő	60031	összeesket	60011
leereszt	60011	összeesküszik	60011
leerősít	60011	összeesküvés	60021
leesik	60010	összeesküvő	60321
leeszik	60010	összeeszik	60010
leevez	60010	összeeszkábál	60011
összeegyezik	60011	padlóbeeresztő	70323
összeegyezett	60011	teendő	10021
összeegyezettethetlen	60031	teleereszt	60011
összeelegyedik	60011	teleesz	60019
összeelegyít	60011	tőkeerős	20033
összeenyvez	60011	zeneegylet	20029
összeereszt	60011	zeneelmélet	20021

## ii (35 rekord)

antiimperialista  
 férfiing  
 hadiipar  
 határkiigazítás  
 háziipar  
 kéziipar  
 kiigazít  
 kiigazodik  
 kiigényel  
 kiiktat  
 kiindít  
 kiindul  
 kiindulás  
 kiindulási  
 kiinduló  
 kiirt  
 kiismer  
 kiismerhetetlen

kiismerszik  
 kiiszik  
 kiizzad  
 kiizzaszt  
 lelkiismeret  
 lelkiismeretes  
 lelkiismeretesség  
 lelkiismeretfurdalás  
 lelkiismereti  
 lelkiismeretlen  
 lelkiismeretlenség  
 lelkiismeretvizsgálat  
 nekiindul  
 nekiiramodik  
 színiiskola  
 színiiskolás  
 tudniillik

- 1 A konferencián szóban rosszul, scserbai teljes stílusban ejtettem az efféle szavakat: itt éppen a gyors, beszélgetési stílus az érdekes, ahol is a két elem a kiejtésben összeolvad. Erre az – ejtésbeli? hangzó stílusbeli? – hibámra Siptár Péter volt szíves felhívni a figyelmemet. Köszönöm. – S ha már az ilyen kritikai megjegyzésekért mondott köszöneteknél tartok: Tarnóczy Tamás a fent ecsetelt szembeállításomat (társadalmi vs. fizikai) túl élesnek találta. "... a fizikai rögzítmények (látható akusztikai lenyomatok) sohasem azonosak a valósággal. A valóság az agyunkban van!" – írta ennek kapcsán magánlevélben. Neki meg külön is köszönöm a hangosan ki sem mondott észrevételt: igyekeztem ennek megfelelően (át)fogalmazni.
- 2 Az olvasó így fogalmat kaphat a DT 2 egészének a felépítéséről. Abban még, az itt közölteken kívül, információk vannak az egyes címszók belső (főleg alaktani) viselkedésére vonatkozóan. A szófajban előforduló nagybetűk – hexadecimális kódok, szintén csak a DT 2-ben.



### Irodalom

Deme László: Hangzásforma és írásforma. Nagy L. János (szerk.): Implom József helyesírási verseny / Rendszer és stilisztika a helyesírásban. Gyula 1992, 25–32.

Hell György: A magyar tőszókincs fonémastatisztikája, szótagtípus statisztikája és mássalhangzókapcsolódásai. Budapest 1983. [Kézirat]

Olaszy Gábor: Elektronikus beszédelőállítás. A magyar beszéd akusztikája és formánsszintézise. Budapest 1989.

# REDUKÁLT FONETIKAI RENDSZER A GÉPI BESZÉDFÖLISMERÉS MEGKÖNNYÍTÉSÉRE

**Tarnóczy Tamás**

MTA Természettudományi Kutatólaboratórium, Budapest

## Bevezetés

Mi a beszédmegértés? Első pillanatban könnyűnek tetszik meghatározni: normális emberi beszéd formájában közölt, hallás útján fölfogott és a fölvevő agyában lejátszódó folyamatok eredményeként megfogalmazódott tartalmú híryanag tudomásul vétele. Engedmények, bővítések, különleges esetek és helyzetek természetesen csökkentik a fenti meghatározás érvényességét, de talán kiindulási alapnak elfogadható. A megértés ellenőrzése például az lehet, hogy a fölfogó személy szóban vagy írásban megismétli a híryanag teljes tartalmát. Itt már több nehézség adódik: pl. nyelvi tudás és az írástudás. A szóbeli ellenőrzés pedig újabb megértési folyamatot kíván meg – ezúttal az eredeti közlőtől –, ami tovább bonyolítja a helyzetet.

Hogyan folyik le a beszédmegértés? Ez nehezebb kérdés, amelynek mai tudásunk szerinti taglalása elengedhetetlen ahhoz, hogy a címben kitűzött feladatunkkal foglalkozhassunk.

Meggondolásainkat két részre osztjuk, korábban csak pszichoakusztikai úton következtettünk arra, hogy milyen folyamatok játszódhatnak le az agyban. Erről mutatunk be néhány eredményt az I. részben. Az utolsó 30 év azonban bámulatosan sok eredménnyel gazdagította az idegfiziológiai és az agytevékenységgel kapcsolatos megértési ismereteinket. Rövid összefoglalásunkban (csak 1985-ig tudtuk az irodalmat nagyjából átnézni), mégis némi áttekintést próbálunk nyújtani a II. részben.

És akkor mi lehet a gépi beszédfölismerés? A meghatározás ezúttal sem látszik nehéznek: normális emberi beszéd formájában közölt, elektroakusztikai eszközökkel fölfogott és rendezett, számítógépi berendezésben alkalmazott folyamatok eredményeként létrejött és az eredeti beszéd minden lényeges elemét tartalmazó híryanag azonosidejű rögzítése. Már nagyjából a kérdésföltevésből is látszik, hogy ezúttal nem valódi megértésről, hanem a géptől megkívánható, lehetőleg szabatos fonetikai fölismerésről van szó. Új követelményként jelentkezik az azonosidejűség, ami az agyműködés esetében magától értetődően teljesül. Ezt a föltételt azonban nem szokás nagyon szigorúan venni; minél több műveletet végez a gép, annál kevésbé lesz pontosan azonosidejű a feldolgozás eredménye. De másodperc nagyságrendű késés már nem engedhető meg. A fölismerés ellenőrzése ezúttal sok gondot jelent. Korrekt követelményként csak az írásbeli rögzí-



tés (beszédre működő írógép fogadható el. Minden más, akár jelzéssel, akár válaszdással hitelesített "fölismerés" csak álbizonyíték, hiszen ezt egy néhány (néhány száz, néhány ezer) elemre begyakoroltatott célgép is megteszi. Ezek után tehát:

Hogyan működik a gépi beszéd-fölismerés? Erre az utolsó 5–10 év fejlődésének rövid áttekintése csak felületes támpontot adhat (ezt kíséreljük meg összefoglalásunk III. részében), azonban a ma már szinte áttekinthetetlenül sűrűn megjelenő irodalmi anyagban minden részletkérdés megtalálható.

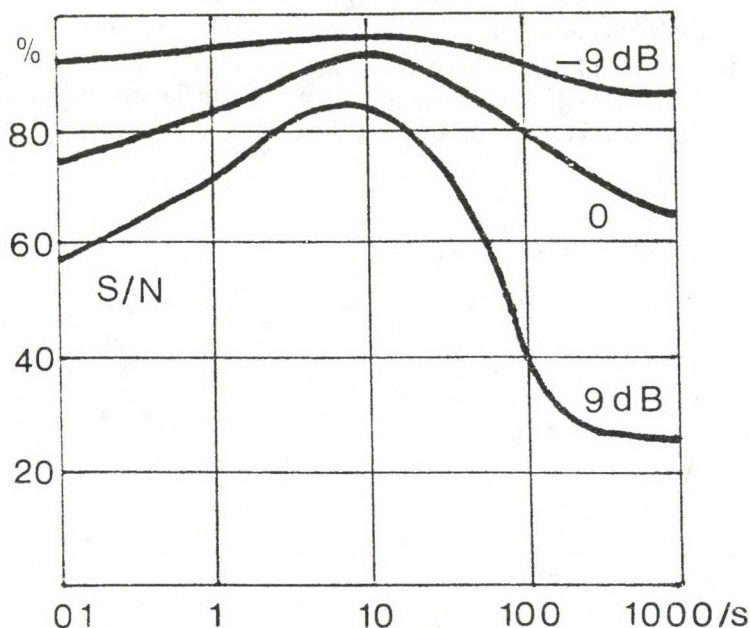
Végül célként kitűzött feladatunkkal is foglalkoznunk kell. Ismerve a gépi fölismerés módszereit és nehézségeit, vajon milyen könnyítést ajánlhatunk fonetikai oldalról a fölismerési követelményünk jobb megvalósítására? Ez a IV. rész vitaanyaga.

## I. A beszédmegértés pszichológiája

A beszédmegértés folyamatának megismerését célzó első kísérletek az érthetőségi vizsgálatok voltak. A harmincas évek óta sok eredményt köszönhetünk ezeknek a mind bonyolultabb körülmények között és mind több paraméter hatását vizsgáló, egyszerű kísérleteknek (Fletcher és Steinberg 1929). Közben lassan – további 30 év alatt – tisztázódtak a legfontosabb kísérleti feltételek, például a vizsgálandó anyag (szótag, szó, érthető, mesterséges összetételű), a hallási körülmények (egyfülű, kétfülű, fejhallgató, teremben végzett, zajjal fedett), végül a lehallgató közösség összetétele (egy-, vagy több személy, kor, képzettség, homogenitás, terhelési állapot) szerint. A kutatandó feladatok a hallásmechanizmus szempontjából értékelhető, sőt a folyamatokra is felvilágosítást adó eredményekhez vezettek.

Példaképpen említhető, hogy a zajjal megszaggatott beszéd érthetősége bizonyos feltételek mellett még erős zaj esetén sem nagyon csökken, ha a megszaggatás fele időtartamú zajközlés mellett kb. 150–200 ms tempóban folyik. Két jelentős kísérletsorozatot érdemes megemlítenünk (Licklider és Miller 1951). Egyik: szaggatott zajjal elfedett szöveg érthetőségének alakulása. A folyamatosan közölt érthetőségvizsgálati szövegek hangzását úgy változtatták meg elfedő zajjal, hogy az fél ideig takarta a jelet, fél ideig szünetelt. Változtatható paraméterek a jelszint-zajszint különbség ( $S/N$ ) értéke, valamint a szaggatás periódusa voltak. Az 1. ábrán az eredeti eredmények lényegét szórás kép nélkül, és a túl sűrű kapcsolgatás értékelhetőségének bizonytalansága miatt csak másodpercenként 1000 kapcsolásig (0,5 ms ideig szöveg és 0,5 ms ideig elfedés) közöljük. Kis intenzitású zaj mellett tehát, ha "fedett" szakaszok gyakorlatilag zajmentesek ( $S/N > 9$  dB), 1...100/s szaggatási frekvencia ( $\Delta t = 1$  s és 10 ms időtartam) között csaknem állandóan 95% marad a szóérthetőség. Ha a kétféle szint közel egyenlő ( $S/N \approx 0$  dB), az ideális periodicitás ( $\Delta t$ ) 50 és 200 ms közé szűkül, de az érthető-

ség megmarad. Ha a zaj 9 dB-lel erősebb a beszéd szintnél, még mindig 85%-os érthetőség tapasztalható. Ez elég meglepő, hiszen ekkora szóérthetőség még 94–97% beszédérthetőséggel egyenértékű, tehát a szaggatott zaj a megértésben nem zavar. Feltűnő azonban az optimális periodicitástól eltérő értékek mellett a rohamos érthetőségcsökkenés. Például az utolsó esetben a 10 ms periodicitás 40%-ra csökkenti az előbbi 85%-os érthetőséget.



1. ábra

Folyamatos (angol) szöveg érthetősége, ha fele időben különböző sűrűséggel szaggatott zörejrel elfedjük.

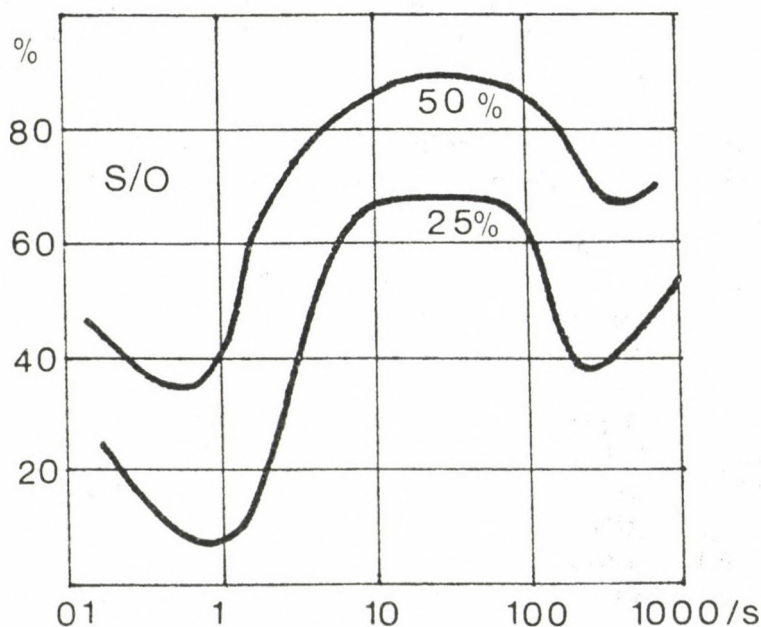
Vízszintes tengely a másodpercenkénti elfedések száma, függőleges tengely a szóérthetőség százaléka, paraméter a jel-zaj arány

Az eredmények javulnak, ha a félidős beosztás helyett relatíve hosszabb beszédidőt engedünk meg, de rohamosan romlanak, ha ugyanezt az elfedő zaj javára módosítjuk.

A 2. ábrán azt az eredményt vettük át, amellyel a szerzők a beszédjel hangtalan állapotokkal való periodikus megszakítását vizsgálták. Ezúttal a paraméter a jelnek a teljes periódusidőhöz viszonyított időtartama. A legjobb érthetőség most a  $\Delta t = 10\text{--}100$  ms periódusidő környékére esik, de rohamosan romlik, ha a szöveg részideje csökken. Ennek gyakorlati jelentősége is van. A 70-es években itthon katonai irányból érdeklődés mutatkozott egy olyan módszer kidolgozására, hogy



hogyan lehetne egyetlen rádióval a beolvasott hírsanyagot különböző hullámhosszakra periodikusan átkapcsolva, tehát olcsó energiával, de mégis kevés érthetőségcsökkenéssel sokfelé továbbítani. A gondolat megoldhatóságát nemcsak az ábra adatai szemléltetik, de az átkapcsolásokkal kényszerűen jelentkező kattánások és egyéb torzítások érthetőségcsökkentő hatására is számítani kell. Két fél-igénybevétellel még talán kivitelezhető, de négy negyed-igénybevétellel már használhatatlan megoldáshoz jutunk.



2. ábra

Megszaggatott (angol) szöveg érthetősége,  
ha csak az idő 50%, ill. 25%-ában halljuk.

Tengelyek, mint az 1. ábrán, paraméter a hallható szövegrész százaléka

Néhány hallástani következtetés ezekből az egyszerű kísérletekből is levonható. Mivel az érthetőségi szövegminták szavak voltak, valószínűsíthető, hogy a 100....200 ms szótaghosszúságnak alapvető szerepe van a megértésben. Egy másik következtetés akkor még nem volt világosan megfogalmazható, hogy ugyanis hallás korrelációs elemzéssel előnyben részesíti a beszédinformációt a zajinformációval szemben. A mai ismeretink szerint a két fül jeltartalma között a hangforrás irányától függően (eltekintve a pontosan szemből érkező hangjel esetétől) jól értékelhető időkülönbség jön létre, ami a homogén jellegű zaj esetén nem jelent információkülönbséget.

Más vizsgálatok a hang irányának kétfülü fölismeréséről azt mutatták ki, hogy a két jelminta közötti 0,01 ms időkülönbséget már érzékeli a fül (Feddersen 1957; cit. Blauert 1974). De mivel ebből az is következne, hogy a jól ismert Gábor-féle formula szerint (Gábor 1946) 10 000 Hz sáv szélességen belül nem tudnánk hangmagasságot megkülönböztetni, nyilvánvaló, hogy a hallás külön hasonlítja össze az időbeli és külön a magasságbeli adatokat. Hiszen szintén közismert kísérleti eredmény, hogy 1000 Hz körül mintegy 3 Hz pontosságú a hangmagasság-érzékelésünk. Gábor Dénes ugyanis az információelméletben egy jel időtartamának ( $\Delta t$ ) és frekvencia-sáv szélességének ( $\Delta \omega$ ) bizonytalanságára a Heisenberg-féle reláció analógiája szerint a  $\Delta t \Delta \omega \approx 1$  összefüggést valószínűsítette.

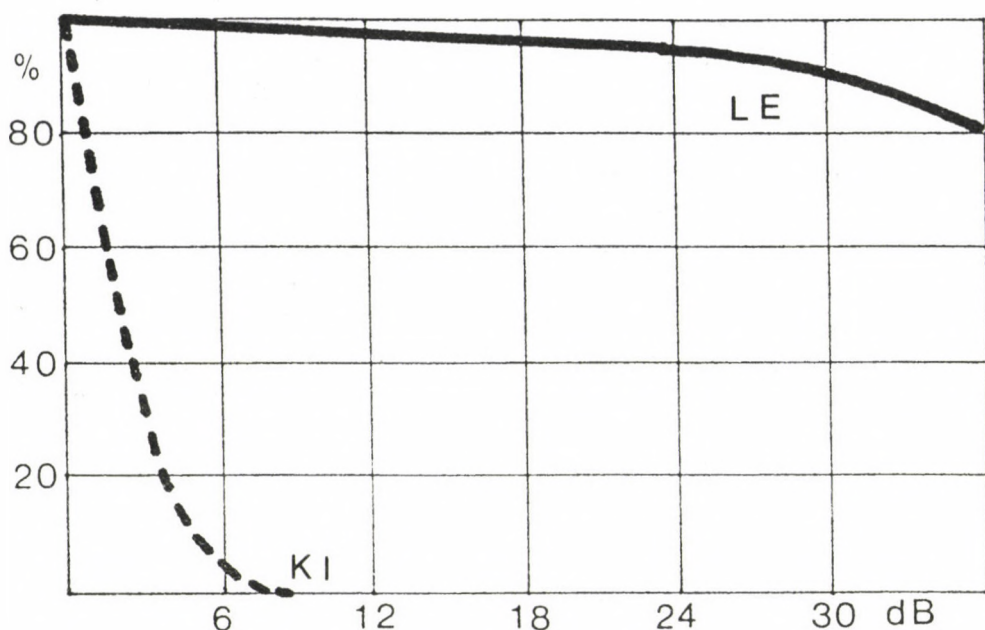
A sávkorlátozott beszéd érthetőségének vizsgálata (Fletcher 1953) arra adott fölvilágosítást, hogy milyen hangoknak, milyen frekvenciatartományban fekszik az érthetőségi csomópontja, illetőleg, hogy milyen tévesztések jelezhetők előre bizonyos frekvenciatartományoknak az átviteltől való kiszűrése esetén. Mindez nemcsak híradástechnikai fontosságú következtetésekre adott lehetőséget, hanem a halláskárosodás helyének érthetőségi vizsgálatokkal való meghatározását is lehetővé tette.

A beszédmegértés rejtett folyamatainak megismerése szempontjából egyik legfontosabb pszichoakusztikai kísérlet volt a Licklider által már 1944-ben végrehajtott, 1946-ban, majd két évvel később munkatársával részletesen közölt (Licklider és Pollack 1948), oszcillogram torzítási technológiával végzett érthetőségi vizsgálatsorozat.

A híradástechnikai megoldás abból állt, hogy a beszédről fölvett oszcillogram csúcsait levágták, majd a maradék jelet többször fölerősítve és újra levágva, végül közönséges négyszögjeleket nyertek. Ezek érthetősége azonban az eredeti oszcillograméhoz képest alig változott: 24 dB levágásig (amikor az eredeti amplitúdónak csak 1/16 résznyi csonkja maradt meg) a szóérthetőség 97% maradt. Tovább folytatva a kísérletet, a jel differenciálásával már csak a 0-átmenetek helyén álló, állandó magasságú impulzusjelek képviselték az eredeti oszcillogramot. Az érthetőség ekkor sem csökkent jelentősen. Ha azonban a kialakult szerkezetet akár csak 0,1 ms-os beosztás szerint "rendezték", az érthetőség azonnal eltűnt.

Az agyi értékelésben tehát nem az oszcillogram formája döntő, hanem a 0-átmenetek rendszere. Erre is megvolt az ellenkísérlet. Ha nem a csúcsokat vágták le, hanem a 0-tengely körüli átmeneti részeket, és pedig csak 2,5 dB értékben (vagyis az átlagos amplitúdó egynegyedét) vágták ki, az érthetőség 30%-ra esett, 6 dB kivágás esetén pedig már csak moraj volt hallható az eredeti szavak helyén. A két meggyőző görbét a 3. ábrán mutatjuk be.





3. ábra

Angol szóérthetőség változása, ha az oszcillogram csúcsait levágjuk (LE), ill. ha a 0-átmeneteket kivágjuk (KI). Vízszintes tengely a csonkítás dB-értéke

A természetes tévesztésekből következtetni lehet a közeli tulajdonságú beszédhangok megkülönböztető jegyeire (distinctive features). Például 50 ms rövidségűre csonkított magánhangzókat sokszor ítélünk mássalhangzóknak. Egy másik lehetőség a beszéderősséggel nagyjából azonos intenzitású, fehérzaj jellelő zavarás ( $S/N \approx 0$  dB) mellett vizsgálni egyes hangok érthetőségét. Példaként zárhangokra adjuk meg az érthetőségi stabilitást angolban és magyarban az 1. táblázatban, a tévesztési indexeket pedig a 2. táblázatban (Miller – Nicely 1955; Tarnóczy 1964).

1. táblázat

Hang	Angol	Magyar
p	50,7	46,5
t	76,6	89,5
k	50,2	78,5
b	75,3	75,0
d	72,7	68,0
g	59,2	66,5

Érthetőségi stabilitás zárhangokra (%)

2. táblázat

Tévesztés	Angol	Magyar
p/t	11,9	0,4
p/k	31,0	0,6
t/p	12,8	23,0
t/k	14,8	10,6
k/p	29,8	4,0
k/t	11,6	3,0

Tévesztési indexek a második hangról az elsőre (%)

Az érthetőségi stabilitás a magyarban általában jobb, a tévesztési indexek egy kivétellel szintén jelentősen előnyösebbek, viszont a magyarban elég sok a zöngés/zöngétlen tévesztés, angolban pedig ilyen nem találtak. Egyetlen vizsgálati sorozat nem döntő, de annyi megállapítható, hogy ebben is vannak nyelvi különbségek.

Az érthetőségi vizsgálatoknak természetesen nagy szórása van. Említettük, hogy az eredmények mennyire függenek a körülményektől. Két dologra utalunk. Az egyik a vizsgálati anyag összeállítása. A folyamatos anyanyelvi szöveg érthetősége – egyébként azonos külső körülmények között – mindig a legnagyobb. Az agyban elraktározott emlékképek, a hiányzó, hibás vagy félreérthető hangelemek utólagos kiigazítása, a szöveg tartalma szerinti várakozás mind fontos kellékei az agyi megértésnek. A legtöbb esetben egy elkezdett szó, sőt egy bizonyos szövegirányban értelmezett mondat is lehetővé teszi a hangfolyamat "előre tudását". A hallgató a közlő egyén nyelvi és nyelvtani hibáit, következetlenségeit az imént vázolt előrejelzés segítségével állandóan korrigálni képes. (Ez a szerencsénk a hadarva, hangkihagyásokkal beszélő politikusoknak és egyéb nyilatkozóknak.)

Van ennek a mechanizmusnak egy további, az agyi megértés szempontjából szintén érdekes következménye. Arra a sokszor megtörténő jelenségre gondolunk, hogy külföldön az utcai zajközegben meghallott idegen nyelvű beszédfoszlányokból nemcsak magyar szavakat, hanem teljes mondatokat is ki lehet hallani. Objektív kísérletek (Warren 1970; Miller et al. 1976) azt tanúsítják, hogy valóságos beszédhangokat elfedő, vagy a helyükbe iktatott – beszédjelleggel egyáltalán nem rendelkező – zörejhangok az éppen odaillő beszédhang érzetét keltik a hallgatóban. Észre sem veszi, hogy valami egészen más ingert kapott, az agya a kívánt beszédhangot illesztette a helyébe. Alaposabb megfontolások arra utalnak, hogy az agyi megfigyelésnek előre be kell állítódnia, hogy beszédjelek (speech sounds) fölismerésére készüljön föl, vagy eleve tudja, hogy zörejeket (non-speech-sounds) fog észlelni. A kettő között – becsapás jellegű – jelsorozat (speech-like sounds) esetében, mint amilyenek a mesterségesen összeállított "beszédhangok", az agy rendszerint jóindulatú "beszédmegértés" állapotra kapcsol.



Ennek a közismert jelenségnek az érthetőségvizsgálatokból való kiküszöbölése érdekében készítették eleinte a "szabványos" összetételű logatomokat postai berendezések érthetőségi ellenőrzésére. Hibái között említhető, hogy ezek nem valóságos hangkapcsolódások, amelyeknek még a nyelvtől való függetlenségét is föltételezték. A kérdés taglalása helyett megemlítjük, hogy magyar érthetőségi vizsgálatok céljára olyan szövegmintákat készítettünk (Tarnóczy 1954; 1961; 1973), amelyek hangstatisztikája, szótagformái és szótagkapcsolatai megfeleltek a magyar szavak általános típuseloszlásának, de – mivel az összeállítás véletlen generálással történt – nem voltunk tekintettel arra, hogy esetenként értelmes vagy értelmetlen szóképek keletkezzenek. Ezzel igyekeztünk megakadályozni, hogy a hallást eleve befolyásoljuk, hogy milyen összetételt célszerű kihallania.

Meggondolásunk helyességét igazolta, hogy még az instrukciók ellenére is visszamaradt a hallgatóban valamilyen kényszer, hogy értelmes hangkapcsolatot értelmezzen. Tömeges érthetőségi vizsgálataink céljára, pl. egy nagy színház megtöltésére, sokszor kaptunk honvédségi segítséget. Humoros példaként említem, hogy az esetek feltűnően nagy százalékában az "arc" szóra adott válasz "harc" volt. Nyilván ez a szókép a kiskatonák agyi emléktárában nagyon fontos.

## II. A beszédmegértés folyamatai

Az emberrel végzett érthetőségi vizsgálatok eredményei csak közvetett módon utalnak a hallás agyi folyamataira. A tudományos kíváncsiság azonban a hallás és az agyi értékelés pontosabb megismerését követeli. Békésy György a belső fül Corti-szervéig bezárólag modellvizsgálatok mellett még tudott emberi kísérleteket is végezni (legalább is cadaveren), közvetlen idegfiziológiai vizsgálatok azonban csak állatokon valósíthatók meg. Az irodalomban közölt adatok jelentős része altatott vagy érzéstelenített macskákon végzett kísérletek eredménye.

Föltételezve, hogy a csigában lefolyó jelenségeket helyesen ismerjük, ezúttal csak néhány újabb eredmény említésére szorítkozunk. Ma már valószínű, hogy a csigában nemcsak durva, hanem idegileg szabályozott, finomabb frekvenciaérzékelés is folyik, ami az agy felé vezető (afferens) és az agyi központokból, vagy felsőbb idegi szintekből a csiga működését szabályozó (efferens) idegpályák együttműködéséből adódik. Újabb vizsgálatok például rámutatnak (Pujol 1990), hogy a 3500 belső szőrsejt közül, ahonnan a 30 000 hallóidegrost legnagyobb hányada, mint radiális afferens indul ki, esetenként egyetlen érzékelő egység is küldhet idegimpulzusokat a felsőbb központok felé. Ez rendkívül érzékeny frekvencia-megkülönböztetést tesz lehetővé. Ugyanakkor az első nyúltagyvelői idegközpont összehasonlíja a bal és jobb fülbe érkező jelek időkülönbségét és már 0,01 ms különbségről (irányérzékelés) is jelentést küld az agynak. Vagyis egyszerre érzékeljük a jel frekvenciájának és időpontjának (pontos) értékét. Eszerint



az agy bizonyos értelemben túlteszi magát a Gábor-féle összefüggésen, tehát közvetve a Heisenberg-féle bizonytalansági reláción is.

A teljes elemzést természetesen a felsőbb idegközpontok és szerkezetek végzik el, hiszen a pontos idő- és frekvenciaérzékelés mellett szinképi és szinképváltozási adatokat és bonyolult összehasonlító (támogatási és gátlási) összefüggéseket is tudomásul veszünk.

Itt egy pillanatra érdemes megállni. Bizonyos eredetű sükettség esetén, amikor az idegvezetés nem sérült, újabban mind gyakrabban alkalmaznak beszéddel vezérelt elektromos ingerlést, melynek rezgésgerjesztőit beoperálják a csigába. Már ötféle frekvenciasávú gerjesztéssel elérhető, hogy a beteg nemcsak a zajokat érzékeli, de kitartó munkával megtanítható a beszédmegértésre is (Cochlear Implants 1985; Ribári 1992). Ebből két következtetéshez juthatunk. Első: a csigában keletkező pontos frekvenciaérzékelés (Pujol és mások szerint nem szükséges hozzá másodlagos idegi szűrés vagy elemzés) lehetővé teszi, hogy a magasabb idegi központok elsősorban a bonyolult időlefolyású akusztikai folyamatokkal (ilyen a beszéd) foglalkozzanak. A második következtetés ennél forradalmibb. Az agy végül is képes arra, hogy nem a megszokott pontos szinképi alakulások helyett, homályosan a lényegre utaló szinképi foszlányokból – "rekonstruálja" a közölt beszéd elemeit és folyamatát. Ha a korábban említett szövegbeli helyettesítő és kiegészítő agyi tevékenységre gondolunk, nyilván nem képtelenség ilyen teljesítményt föltételeznünk. Hivatkozhatunk a "missing stimulus events" jelenségre is, amit az optikából jól ismerünk. Az implantációnál azonban nemcsak erről van szó, hanem többek közt a beszéd szinképi elemeinek átranszformálásáról bizonyos olyan frekvenciasávokra, amelyek akusztikai jelentését az operált személy későbbi gyakorlatában megtanulja beszédhangokként értelmezni. A gerjesztő beépítésével "rehabilitált" nagyothallók a gyakorlatban elérik a beszédmegértés kisebb-nagyobb fokát.

A rezgések formája, frekvenciája és erőssége az agy különleges kapcsolatai révén nemcsak a hallópályán át juthat el az agyi halló- és érzékelő központokba. Például ujjakon keresztül kapott (taktilis) rezgések is értékelhetők beszédhangok gyanánt (speech-like vibrations). A tíz ujj rezgési ingerlésének érzékelését is meg lehet tanulni beszédjelekként értelmezni.

További megfontolások valószínűsíthetik, hogy az agyi feldolgozás a már említett 0-átmeneti statisztika mellett időben folyamatosan Fourier-elemzésre is támaszkodik, hiszen a formánsok azonosidejű fölismerése erre mutat. Ezeken fölül még mintegy 6–8-féle matematikai munkát, pl. korrelációs elemzést, konvolúciós integrálást, fordított energiaszinkép (cepstrum) számítást, továbbá statisztikázást kell az agynak végeznie, hogy valóban – legalább fonetikailag pontosan – értse a beérkező jelfolyamat adatsorát. Ismét csak egyetlen, eddig kevés figyelemre méltatott adat jelentőségére kívánok rámutatni. Elfogadott tény, hogy a hangjel hatására a csiga belső csatornájának folyadékában nyomáshullám fut vé-



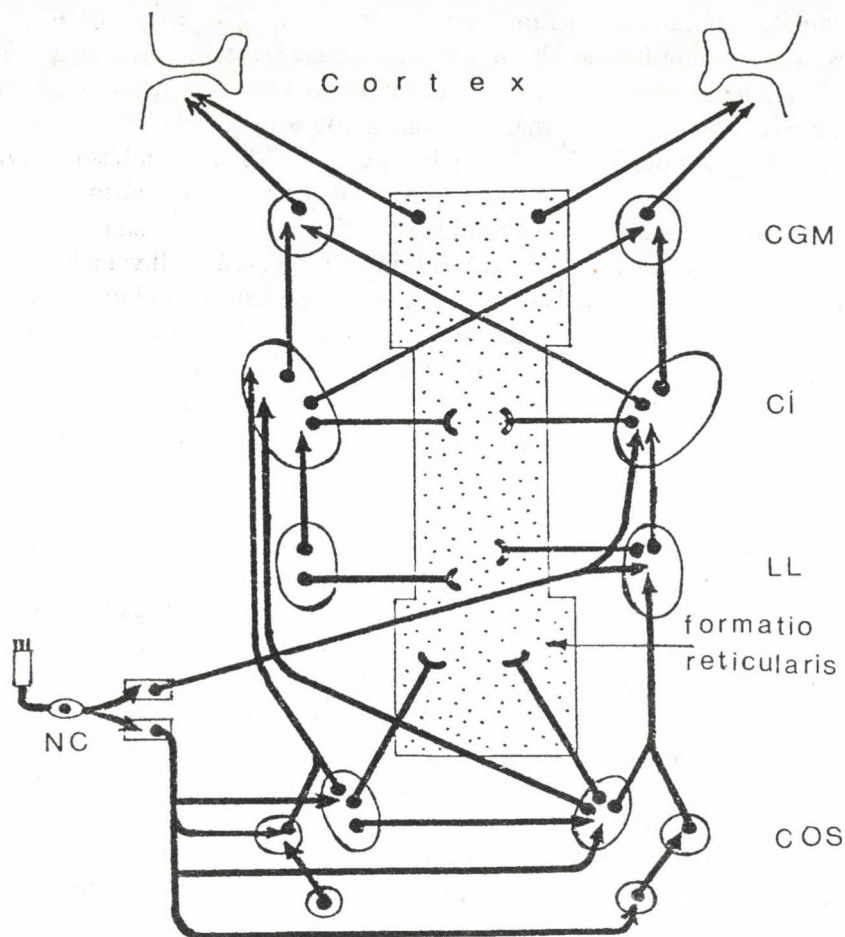
gig, amely tehát nem pontosan ugyanabban az időpillanatban ingerli az alaphártya egész területét. Békésy eddig meg nem cáfolt adatai szerint (Békésy 1943) a csiga alapjától számított kb. 15 mm hosszú szakaszon az ingerlés 0,01-től 0,1 ms időt késik. Az utóbbi helyen kb. az 1500 Hz-es érzékelési terület található. Tovább haladva az alaphártyán, 25 mm-nél (500 Hz) a késés 1 ms, a csiga csúcsához közeledve (50 Hz) pedig már alig vánszorog a nyomáshullám, és késése 5 ms körüli. Vagyis a beérkező komplex hangjel időben szétszakad és az érzékelő sejtek először a színekép magas hangelemeit továbbítják (pl. a formánsokat) és csak sokkal később – 3-4 ms nagyon hosszú idő – a mély hangokat (pl. az alaphangot). Ezt az agyi értékelés leírásában szintén figyelembe kell venni.

A VIII. agyideg a hallópálya és az egyensúlyérzékelő pálya idegfonalait egy kötegben vezeti el. De egymásra nincsenek hatással. A hallóidegek axonjainak egyik sajátága, hogy elektromos szigeteléssel vannak ellátva.

Tekintsük meg a nagyon vázlatosan összeállított 4. ábrát! A valóság természetesen sokkal bonyolultabb, és egyszerűsítésül csak a bal fül idegi kapcsolatait rajzoltuk be. A másik oldal ehhez képest szimmetrikusan viselkedik, ezt mindig vegyük figyelembe.

Egyes idegrostok elemi tevékenységét már régen elég pontosan megvizsgálták, valamint jól ismerjük a gerinces állatok hallóideg kapcsolatainak menetét is, ami nagy valószínűséggel azonos az emberével. Alapos kísérleteket elsősorban macskákon végeztek, de az eredményeket főemlősökön is hitelesítették. A fő nehézség azoknak a jellegeknek (features) a megtalálása, amelyekre az egyes neuronok érzékenyek. Hallás szempontjából alacsony szinten tisztázódni látszik az oldalirányú gátlás szerepe a színeképi érzékelésben. Egészen biztos, hogy a hátsó csigamag erősen reagál akár az amplitúdó, akár a frekvencia modulációjára. Azonban nagyon nehéz ilyen jellegeket találni magasabb szinten, főként az agykéregben, ahol merőlegesen sok egymás alatti neuron azonos viselkedést mutat (kolumnáris szerkezet), viszont a tangenciális kapcsolatok áttekinthetetlenül sok irányúak. Az agykérgi kapcsolatokról később lesz szó, de biztos, hogy a szervezetben a megoldás inkább egy hologramos rekonstrukcióra hasonlít, mint egy fényképére, vagyis a végső megismerés kialakulásához valószínűleg nem pont-ról-pontra következtetéssel, hanem sok elemi információs adat valamilyen különleges integrációjával jutunk el. Erről azonban többnyire csak föltételezéseink vannak.

A 4. ábrával kapcsolatban csak néhány lényegesebb megjegyzést teszünk, pontosabb ismertetés fiziológiai könyvekben bőven található (Schmidt 1980; Forsmann – Heym 1982; Pickles 1982; Jones – Peters 1985).



4. ábra

A hallóideg pálya erősen vázlatos "huzalozási" rajza a bal fültől az agykéregig. NC (nucleus cochlearis sinister): a bal oldali csigamag, COS (complexus olivae superioris): bonyolult átkapcsoló központ közvetlenül az agyideg belépése helyén az olivában, LL (lemniscus laterális): a második idegi átkapcsolódás helye, CI (colliculus inferior): a harmadik átkapcsolódás, CGM (corpus geniculátum mediale): a hallópálya idegrostjainak utolsó átkapcsolódása a thalamus hátsó pólusánál, ahonnan az axonok a homloklebenyek hallásfeldolgozó területei felé sugárzódnak. A pálya teljes hossza 5-6 cm



1. A csigából kivezető idegköteg az oliva helyén fut be a hátsó agyba. Innen 4-5 neuron átkapcsolásával jut mind följebb, egészen az agykéregig (cortex), ahol az idegi információk együttes hatására a hangérzet különféle összetevői keletkeznek. Az egyes átkapcsoló állomások bizonyos neuron központokban találhatók (az ábrán megnevezve), ahol az axonok szinaptikus érintkezéssel adják tovább az elektromos impulzusokat a következő idegnek.

2. Az átkapcsoló állomásokon nemcsak afferens, hanem efferens jelzések is továbbítódnak (ez nem látható), továbbá elágazások és ellenoldali átvezetések fordulnak elő. A keresztezés mintegy 80-85%-os, de ezeket is csak főbb elágazásaikban tüntettük föl (afferens: agykéreg felé, efferens: ezzel szembe futó ideg).

3. Aránylag legjobban ismertek a csigamag (nucleus cochlearis) és a közvetlenül mellette, kissé lejjebb található oliva superior terület neuronkapcsolódásai. A csigamag hátsó (dorsalis) csoportja különösen érzékeny a hang-modulációkra, amellet a következő lépésben kikerüli az olivát és egyenesen a következő állomásokhoz (lemniscus lateralis, ill. colliculus inferior) kapcsolódik. Az oliva különböző átkapcsoló területeiről az a különös, hogy mindkét fül információit feldolgozzák és különbség képzéssel, többek közt a két fül közötti időkésést – a már említett, 0,01-0,1 ms közötti pontossággal – érzékelik és tovább jelentik.

4. Kevés neuronról tudjuk, hogy pontosan mire érzékeny, de az biztos, hogy a hangmagasság érzékelése, az ún. hangolási görbe, valahol a harmadik-negyedik állomásnál a legélesebb. A magasabb központokban egyébként mind összetettebb és bonyolultabb tulajdonságok tudomásul vétele folyik, például a hangkezdés és befejezés formai érzékelése, az időbeli színekpi változások, vagy a színekpi összetettség fölismerése. Mérésekre alapozott megfigyelés, hogy a magas hangokra érzékeny idegrostok a mélyekre (pl. a magánhangzókra) azonosan reagálnak, vagyis az állandósult állapotú ingerhatás nem vált ki bennük semmi változást, azonban a felnövekvést és az elhalást (a hang elejét és végét) időbeli változásként tudomásul veszik. Maga a színekpi forma az agyban több helyen is megjelenik, mint az alaphártya gerjesztési helyzetének helyi megismétlése (tonotop leképzés). Ebben ismét jelentkeznie kell a színekpi időben fölülről lefelé megjelenő kialakulásnak.

5. A másik oldali (contralateralis) idegpályákhoz való átkapcsolódás több következménnyel jár. Ezek idegvezetékei általában vastagabbak, tehát gyorsabban és nagyobb intenzitással továbbítják a jeleket, így azok előbb jelennek meg egy magasabb központban, mint a saját oldali (ipsilateralis) információk. A jobb fül fölötti jelei tehát hamarabb (a becslés 8 ms), de mindenesetre nagyobb intenzitással jutnak el a bal, mint a jobb hallásmezőbe.

6. Az egyes szinaptikus átkapcsolódások természetesen idővesztést jelentenek. Egymás utáni kapcsolódás esetén a késés akár 5-10 ms-ot is elérhet. Ha ehhez még az idegimpulzusok terjedési idejét is hozzávesszük, az előbbi adat kétszeresére számíthatunk. Csakhogy addigra a közbenső állomásokon sok min-



den tisztázódott és a jelentések tömege fut be az elsőrendű hallásmezőkhöz. Az információs adatok továbbításának gyorsítását szolgáló említett lehetőségek mellett a növekvő számú párhuzamos szinapszisok kisebb ellenállását is elképzelhetjük.

7. Megjegyezhető még, hogy az ábrán is jelölt hálózatos szerkezet nem információ továbbítására szolgál, hanem az érzékelő agykéreg "riasztására", hogy a hírek főlvetelére és megértésére minél hamarabb fölkészüljön (ez többek közt a vérellátás fokozódásában is jelentkezik).

Röviden azzal is foglalkoznunk kell, hogy az agykéregben mi történhet. Az állatkísérletek mellett az agy kutatásában új ismereteket hozott a betegeken is végezhető radioaktív nyomjelzés módszere, a számítógépes rétegvizsgálat, a mágnes encephalográfia és a pozitron emissziós tomográfia. Ezen módszerekkel lehetett megállapítani a hallás- és beszédközpontok helyét, sőt egyes beszédképzési és -értékelési tevékenységek pontos területeit, vagy pl. azt, hogy a bal félteke Wernicke-mezőjének 10%-kal erősebb a vérellátása, továbbá, hogy nem beszédjellegű hangok a jobb oldali féltekét terhelik jobban.

Mai becslések szerint az ember 40-60 milliárd neuronjának nagyobbik fele az agyban koncentrálódik, annak is főleg a kérgi szürke állományában. Az agy tevékenysége szervi információk földolgozása mellett elsősorban aktív tevékenység: a külső világ hatásainak rendszerezése, analógiák fölismerése, a logikus gondolkodás és az alkotó munka. Mindezt a sokmilliárd neuron bonyolult kapcsolódásai teszik lehetővé.

Az agy különleges felépítése folytán igen nagy külső felülettel rendelkezik; itt helyezkednek el az értékelő neuronok. Ez a szürke állomány a hallás és beszéd központokban 2-3 mm vastag, és benne merőlegesen lefelé 6 "emelet" mélységig különböző neuronszerkezetű rétegek találhatók (isocortex). Állatkísérletek szerint merőleges irányban egymás alatt általában hasonló tulajdonságú idegmagok rendeződnek, ezek egymást segítik, a félreértést helyrehozzák, az élettellel együttjáró sejtpusztulást pótolják. Az egymás mellett, akár egészen közel található neuronok viszont sokszor teljesen eltérő viselkedésűek.

Megállapították azt is, hogy az agy igen széles – egészen más feladattal megbízott – területei is összeköttetésben vannak a beszédelfogó és értékelő területekkel, és pedig szinte közvetlenül vagy kevés átkapcsoló lépéssel. Ez két dologra utal: 1. hogy a beszédmegértés igen sok mellékes dologtól is függ és 2. hogy ez a befolyásolás rendkívül gyorsan megy végbe (Seldon 1985). Több kísérlet igazolja, hogy akár a szintaktikus memória hirtelen ébresztése, vagy asszociációk fölismerése mennyire segíti a megértést. Az eredmény természetesen az egyén nyelvismeretétől, műveltségétől és intelligenciájának fokától (IQ) függ.

Az agy két féltekéje nagyjából szimmetrikusan látja el feladatát, az egymásnak megfelelő kétoldali felületek egymással idegi összeköttetésben állnak. Az agyvelő mégsem teljesen szimmetrikusan működik. Ez az aszimmetria éppen el-

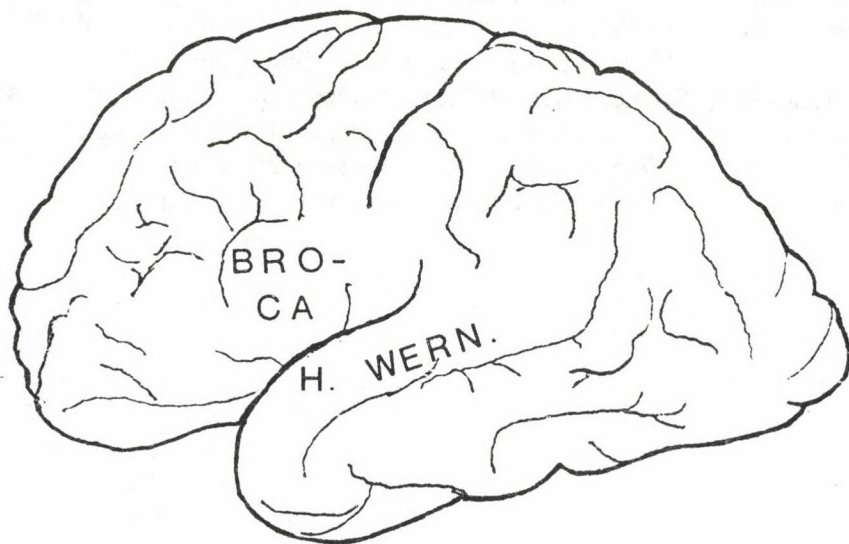


sősorban a beszédre, illetőleg a zenére vonatkozik. Csimpánznál még találták némi nyomát az aszimmetriának, de alacsonyabb rendű majmoktól lefelé már teljes a szimmetria (Seldon 1985). Ez az agyi sajátság tehát jellegzetesen emberi tulajdonság, ahogy a beszéd is az. A bal-félteke a beszédképzés és megértés tekintetében domináns, és egyben gondolkodásban is vezető szerepű oldal. Viszont a jobb félteke feldolgozási folyamata inkább a zenei jellegű érzékelést részesíti előnyben (Kimura 1961; 1964). Sőt, mivel az idegvezetés a bal féltekébe a jobb fülből határozottabb, a jobb fül a szöveg megértésben előnyösebb helyzetben van (Kimura 1967). Számos kísérlettel igazolták azóta, hogy minél több – akár színképi, akár időbeli – változás van egy folyamatos híryananyagban, annál inkább áttolódik a megértési előny a jobb fülbe, vagyis közvetve a beszéd megértő bal féltekébe. Valószínűleg ez az oka annak is, hogy a monoton hanghordozású vocoderek szövege azonnal érthetőbbé válik, ha a hangmagasságot akár szabálytalan összevisszaságban is változtatgatjuk.

A beszéd megértéshez természetesen a fonémák színképi alakzata is hozzájárul, ami jobb oldali földolgozást igényel. De mégsem ez a lényeg, hiszen a színkép időbeli változását ismét bal oldali tevékenység ellenőrzi. Az általános vélemény tehát abban egyezik meg, hogy a mássalhangzókat inkább a bal oldal, a magánhangzók állandósult állapotát pedig a jobb oldal dolgozza föl, de a kapcsolat közöttük gyors és állandó. Az összeköttetés a corpus callosumon keresztül a két félteke között szinte pontról pontra megvan. A két oldal tehát kisegíti egymást. A vokálisok információs adatai a színkép állandósult részének erősen zenei jellege miatt a jobb oldalra is befutnak, a tranzienst folyamatok, valamint a 40-50 ms időtartamú mássalhangzók feldolgozása azonban a bal félteke feladata. A jobb oldali Wernicke-mező tehát segíti a végső megértést, és mindenképpen innen várhatjuk a beszélő egyéni hangszínének fölismerését. Sok megfigyelés utal arra is, hogy a cortex alatti neuronális szint a thalamusban már elvégzi ezt a kétfelé osztást és hogy a rövid idejű emlékezés is itt alakul ki (Ojemann 1983).

Az 5. ábrán az ember bal oldaláról nézett külső agyfelületének vázlatos rajza látható, amelybe az akusztikailag fontos mezőket is bejelöltük. Ezek az említett korszerű vizsgálati módszerekkel egészen pontosan meghatározhatók. A primer terület a H helyen van (hallóközpont), az ide beérkező adatok oszlanak el vagy találkoznak a máshova befutott információkkal. A hallóközpont a két féltekén szimmetrikus szerepű. Ugyancsak szimmetrikus a Wernicke-mező elhelyezkedése is, azonban a beszéd megértés mégis a mindenhol a bal féltekébe érkező jelentések feldolgozása alapján jön létre. A nagyobb idegi tevékenység nagyobb vérellátást is igényel, és ezt a bal oldali Wernicke-mezőre már biztosan kimutatták. Sérülése beszédzavart (Wernicke-aphasia), sőt emlékezet és gondolkodás kiesését is eredményezhet. A motoros beszédközpont, a Broca-mező aszimmetrikusan fejlett: a jobbkezes embereknél csak a bal oldalon található (Forssmann-Heym 1982). Sérülése teljes beszédképtelenséggel járhat. A beszéd megértés és a

logikus gondolkodás képessége azonban ilyenkor nem sérül. Ezzel a ténnyel is cáfolható a régebben divatos motoros beszédmegértés elmélete, vagyis hogy a beszédmegértéshez a hangképzés utánérzése vagy egyenesen utánzása hozzájárulna.



5. ábra

Az emberi agy bal féltékéje bal oldalról szemlélve.

Vázlatos rajz a hallóközpont (H), valamint a motoros (BROCA) és a földolgozó-megértő (Wernicke) beszédközpontok körülbelüli helyével

Azonban az eddig leírt aszimmetria nem mindenkire igaz! Az emberek legnagyobb része valóban jobbkezes és motoros beszédközpontja a bal féltéken van. A balkezeseknél a beszédközpont valószínűleg átfordul a jobb féltékére. A beszédmegértés azonban 97%-ban a bal féltékéhez kötött, mégis csak 65%-ban mutatták ki, hogy ez a félteke a fejlettebb. A bal oldali fejlettség jelenti a gondolkodás magas fokát. A kimondottan nagy zenei tehetség viszont kapcsolatban van a jobb félteke fejlettségével, ami a statisztika szerint az emberiség 11%-ának sajátja (Kimura 1961). A két dolog nem választható élesen széjjel. Mindenki érti a beszédet akkor is, ha a közlő történetesen balkezes, bár a dadogás (Broca-mező) és a balkezesség kapcsolta kimutatható. Ellenpélda, hogy a balkezesek erőszakos átszoktatása jobbkezességre, beszédzavart okoz. Ugyancsak téves volna arra gondolni, hogy minden jó zenész balkezes.

Az összefüggés és elválasztás az egyes területek között szintén nem egyértelmű. Leírtak olyan esetet, amikor valaki teljesen elvesztette beszédképességét, de



ugyanakkor tökéletesen megértette a nyelvet teljes grammatikájával együtt (Lenneberg 1962).

A dolgok pontosabb tanulmányozása alapján ma már ismerünk olyan eseteket is, amikor a sérült bal félteke szerepét átvette a jobb. Ugyanis a két félteke között rendkívül sok az átkapcsolás. A Nemzetközi Pszichológiai Kongresszuson Brüsszelben (1992) Kimura asszony előadásában úgy foglalta össze, hogy a hal-lószervi ideginformációk megjárják mindkét féltekét és az összegzett híryanag-ból készíti el a bal félteke a végleges beszédmegértést.

Mindezekből az ismeretekből kevés támpontunk van a megértés folyamatának modellezésére. Talán ebben segíthet az idegi kapcsolatok tanulmányozása. Vannak ugyanis valószínű magyarázatok az adatrögzítés idegi módosításaira. A rögzített adatok (tudattalan emlékképek, összehasonlításra lehetőséget adó agyi "ikonok") idegi biokémiai-elektromos úton jönnek létre.

A rövid idejű emlékezés például elképzelhető úgy, hogy egyes neuroncsoportok között körbefutó zárt lánc alakul ki. Az ilyen láncok külső ingerlés nélkül öngerjesztő hatás révén információtároló jellegűek (Raschewsky 1938; cit. Ádám 1975). Mi történik ugyanis ilyenkor? A tartós igénybevétel a természetben mindig erősítő hatású. Ez esetben föltételezzük, hogy a sokszoros működtetés növeli a szinapszisok összfelületét, akár úgy is, hogy újabb szinapszisok létesülnek. Egyébként valami ilyesmi a "tanulás" folyamata. Még az is megjegyezhető, hogy minden újabb szinaptikus felületet növeli a posztszinaptikus potenciált, vagyis csökkenti az ingerküszöböt.

Egy másik fontos – de hosszú idejű – emlékezési folyamat a fehérjeszintézis, amely az idegmagok sejttálmányán belül megy végbe. Maga a neuron a testben a legintenzívebb RNS (ribonukleinsav, amelyből fehérje alakul) termelő. Az RNS anyag a citoplazmában összpontosul, a fehérjeátalakulás a riboszómákban történik. Érdekes, hogy ha például 20 éves korban 1 ng/sejt (1 nanogram= $10^{-9}$  g) az RNS tartalom, ez a 40-es életévekben duplájára növekszik, majd a 80 év körül visszaesik a kezdeti értékre. Ez jelenti a megszerzett ismeretek egy részének elfelejtését. Ha az ember információfelvétele 25 bit/s, napi 10 órával számolva, 10 év alatt akár  $4 \cdot 10^9$  bit információt képes rögzíteni. Az ehhez szükséges RNS mennyiség már néhány neuronban benne van (Ádám 1975). Persze a rögzítés később törlődik, ha nincs megfelelő gyakorlat (pl. egy nyelv elfelejtése).

### III. Gépi beszédfölismerés

Az eddig áttekintett eredmények néhány következtetést lehetővé tettek a beszédhangok fölismerése és a beszéd megértése területén, de még nem vezettek el odáig, hogy a megértés agyi stratégiáját megismerjük. Annyi biztos, hogy az egyedülálló beszédhangok is és a szótagok is fölismerhetők, tehát a fölismeréshez nem szükséges ún. kontextus, vagyis teljes értelmű szó. Bizonyára van a



nyelvi kategóriákat megértő magasabb fokozat is az agyban, ahol a beszédhangok fonémákká formálódnak, szótanilag és mondattanilag megértett szöveggé rendeződnek össze. Ebben a rendeződésben kapnak szerepet azok a szintaktikai és szemantikai javító és kiegészítő tevékenységek, amelyek a szövegmegértést szolgálják. Ez azonban már nem egyszerű fölismerés, hanem a nyelv teljes ismeretét föltételező rendező és megértő tevékenység. A szövegmegértés mindig tökéletesebb (jobb százalékarányú), mint a szómegértés, ez utóbbi pedig jóval eredményesebb, mint a szótagfölismerés. A legtöbb hibával jár a hangfölismerés, pontosabban annak fonemikus azonosítása.

Nagyjából a mesterséges fölismerő szerkezeteknek is hasonlóképpen kellene dolgozniuk. Valamikor úgy képzelték, hogy a közönséges Fourier-elemzés maga is megoldja a feladatot. Ennek cáfolata az 1960-as években már világosan megfogalmazódott (Tarnóczy 1965). Érdekes azonban, hogy egyes kutatók, akik abból indulnak ki, hogy az akusztikai jelenség a beszéd minden lényeges elemét tartalmazza, később is megkísérelték a színekép időbeli változásának pontos elemzésével megoldani a mesterséges fölismerés feladatát (Stevens 1980). Kísérleti adatok azonban arra mutatnak, hogy a természetes beszédmegértésben fonetikai szinten fölismerhetők olyan töredékes elemek, amelyeket a legtökéletesebben gördülő színeképek vagy oszcillogramok sem mutatnak ki (Tarnóczy 1987). A gépi beszédfölismerés tehát más úton fejlődött és azóta számos szép eredményt ért el. A végleges megoldáshoz mégsem tudott eljutni, mert más módszereket alkalmaz, mint az idegi-agyi minta.

A drámaian rohamos fejlődés néhány technikai adata röviden a következő.

**1946:** a folyamatos színeképváltozásról (Visible speech) leolvasható szöveg (ezt nem a gép ismeri föl, hanem a kép látása alapján az agyunk), **1958:** az első digitális számítógépek, **1962:** impulzus kód moduláció (PCM), **1965:** a gyors Fourier-transzformáció (FFT), közelítés az azonosidejű megoldásokhoz, zajokra egészen jól bevált, **1970:** integrált áramkörök, digitális tárolás és feldolgozás, áramköri algoritmusok, **1975:** a rejtett Markov-lánc (HMM) bevezetése, **1976:** először foglalkoznak az idegi kapcsolatok matematikájával, **1981:** az első elektronikus morzsa (chip) elkészítése, **1985:** a morzsák rendszeres fölhasználása az elemző és fölismerő szerkezetekben (az 1,8  $\mu$ -os technológiával 1 utasítás földolgozási ideje 200 ns, **1986:** az idegi hálózatok elektronikus utánzása (perceptronok), **1989:** 500-szavas szótáranyag fölismerése betanítása (1  $\mu$ -os technológiával a földolgozás üteme 10 millió utasítás-másodperc), az elektronikus idegi hálózatok (TDNN=time delay neural networks) előretörése a kutatásban, **1991:** laboratóriumban elkészült, beszéd-től függő 5000-szavas fölismerő gép, fejlesztés a 0,35  $\mu$ -os technológia felé.

Az áttekintés valóban lenyűgöző, 10 év alatt az 1/s számítási teljesítménytől a 25 millió/s-ig jutottunk el. Ámde azt is meg kell vizsgálni, mire képesek most a fölismerő gépek. Egy-egy részfeladat elvégzésében az agyi teljesítmény 1 %-ét



már elérik, de hiába szárnyalja túl egy algoritmus számítási sebessége az agy hasonló képességét, egyetlen gép nem képes egyszerre párhuzamosan 10-12-féle, vagy még több számítási és földolgozási munkát végezni, és ennek alapján összegezni a fölismerését. Az agy viszont nemcsak a mennyiségekkel bíbelődik, hanem – egyelőre ismeretlen, de igen bonyolult módon – magát a kvalitást érzékeli és veszi tudomásul. Nem szokták külön hangsúlyozni, de napi tapasztalat, hogy a zenei hallással szemben a magánhangzók fölismerésére mindenkinek "abszolút hallása" van.

A továbbiakban röviden azt kell áttekintenünk, mi a gépi fölismerés módszere. Az akusztikai folyamatot a mai berendezések tökéletesen át tudják alakítani digitális formába, ebben azonban alig vannak állandósult (stacionárius) részek, az átmeneti (tranzienst) szakaszok túltengenek. Már az 1930-as években tudták (Gemelli 1934), hogy a legállandóbb szakaszok, a magánhangzók oszcillografikus képei, de ezek is csak mintegy 30-40%-ban tartalmaznak ún. tipikus (vagyis azonos alakú és összetételű) periódusokat, a többi (bevezető, átvezető vagy befejező) szakasz folytonosan változik. Az első nehézség tehát abból adódik, hogy az egyes hangjellegek között ki kell jelölni az elválasztás helyét (szegmentálás). A mai állapotban ezt a nehéz matematikai feladatot általában 10 ms terjedelmű szakaszok elemzésével, például Hamming-ablakos gyors Fourier-transzformációval, vagy még inkább fordított energiaszínkép (cepstrum) számítással oldják meg. Az eljárás szellemes és gyors, de az ablak formája rossz. A rövididejű agyi emlékezet ugyanis nem szimmetrikus, hanem 3-10 ms a fölnövekedési ideje, de 150-200 ms múlva hal el, tehát a múltba nézve lassan csökkenő emlékezéssel (ahogy a neuronok "kifáradnak") veszi tudomásul és integrálja a hangfolyamatokat. Ezáltal jelennek meg a tudatban, és választódnak jól el egymástól a "gondolati fonémák". Ezen adatok birtokában tudjuk a fonémák sorát, amelyek végül is a szavakat és mondatokat kitöltik, írásjelekkel idejében rögzíteni.

A beszéd fölismerő gépek a feladat jellege szerint különféle stratégiával készülnek. Ez már a raktározandó és azonosítandó anyag kiválasztásakor is jelentkezik. Egy 5000 irodai szót vagy kifejezést fölismerő, laboratóriumban (IBM) működő szerkezet szótárkészletét üzleti szövegek és levelezés 27 millió (!) szavából válogatták ki és állították össze. Még fontosabb annak a célnak a meghatározása, hogy milyen elemek fölismerését akarjuk megoldani. Vizsgálhatunk hangjelleget, fonémát, elkülönítő jegyeket, esetleg allofónokat, morfémaakat, szótagokat, szavakat vagy egész mondatokat. Az értelem nélküli hangelemek vizsgálatában figyelembe kell venni, hogy az ún. kontextus, vagyis más hangelem közvetlen kapcsolata nem okvetlenül befolyásolja a fölismerést. Közismert továbbá, hogy az előre ismert téma (pl üzlet, jegyváltás, úrutazás) jobb fölismerést eredményez, és hogy a szófelismerés vagy a mondatfölismerés más-más eljárást követel, s egyben fokozatosan egyszerűbb is.



Egy másik fontos stratégiai kérdés a már említett elválasztási föladatra való fölkészülés. Ma egy  $1\text{ cm}^2$  felületű  $1,2\text{ }\mu\text{-os}$  elektronikus morzsa 300 000 tranzisztort tartalmaz, és ennek megfelelő sebességgel képes dolgozni, de nyilván más algoritmusok kellenek a CVCV típusú (pl. japáni), mint a CCVVCCV típusú (pl. angol) szavak fonémákra szeletelésére. A szétválasztási stratégia tehát nyelvfüggő.

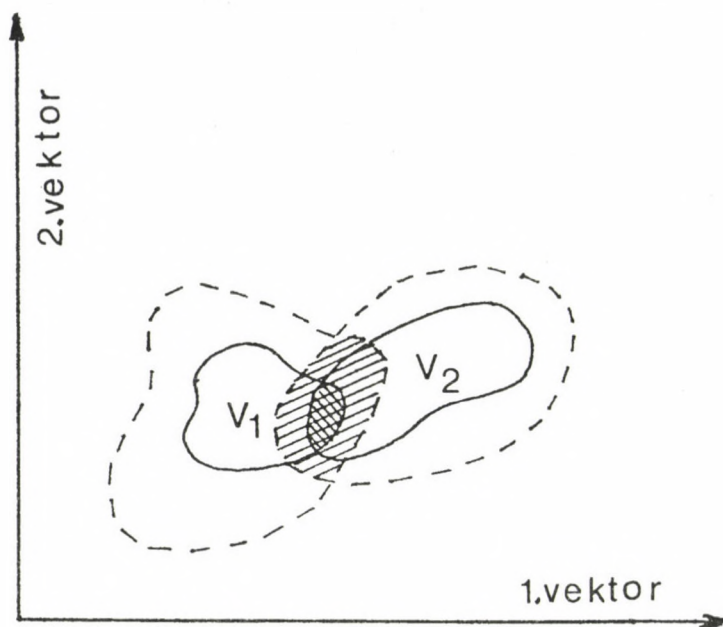
Ezzel a nehézséggel kapcsolatban említsünk meg egy érdekes összehasonlítást. Bár az agyi neuronok számával és eloszlásával kapcsolatban elég bizonytalanok az adataink, szövettani mikrofölvételekből (Seldon 1986) becsülni lehet a beszédmegértésben résztvevő elemek számának legalább a nagyságrendjét. Eszerint  $1\text{ cm}^2$  felület alatt egymillió oszlopban helyezkednek el a neuronok, ezen belül az isocortex hat rétegében legalább ezer neuron és idegi kapcsolat van, vagyis minden  $\text{cm}^2$  felület 100 millió adatot szolgáltat a beszédmegértéshez. Ezért tud egyetlen ember akár 8-10 nyelvet megtanulni.

A harmadik – talán a legnehezebb – kérdés megoldása a kiejtési különbségek összezésizolása és a közös lényeges információs elem biztos fölismerhetősége, ami nemcsak a géptől, de a felhasználó személytől is nagy alkalmazkodó képességet tételez föl.

Egyetlen személy ejtésére alapozott fölismerő szerkezet betanításakor lényegében a tulajdonos tanul hozzá a géphez. Megszokja a pontos kiejtésű tagolt beszédet (a hangsúlyt és az időtartamokat a gép "vetemíti"), szükség esetén megismétli a szöveget. Ilyen szerkezeteket általában műhelyben, irodában használnak valamilyen emlékeztető vagy a jelszóval előhívható anyag (esetleg szóbeli válasz) megjelenítésére. A beszédsebességnek azért korlátai vannak. A természetes beszédmegértés sem megy hadaró beszédközlés esetén. A 12-15 hangelem/s beszédsebesség motorikusan ugyan könnyen túlszárnyalható (a televíziós közvetítők hanghordozása), de a megértéshez erős figyelem kell, ezért a hallgató hamarabb fárad ki, mint a beszélő. Ezt az ismert hatást az egyetemi előadások szósebessége és a megértési százalékok közötti összehasonlítás a gyakorlatban jól szemlélteti. A gépek teljesítményének jóval szerényebb lehetőségei miatt tehát pontosan tagolt és nem túl gyors beszéddel kell a "betanításhoz" fogni.

A szolgáltatás rendeltetésű gépek személytől független fölismeréssel dolgoznak. Ilyenkor jóval szűkebb szótáranyagot sok személy kiejtésének középértékelésével kell a gép tárolójában rögzíteni. A legtöbb ilyen gép jegyváltásra, repülőgép rendelésre, művelet szabályozásra készült, rendszerint azonos szöveg többszöri bemondásra lép működésbe, sőt szükség esetén gondoskodni kell emberi beavatkozásról is. A 6. ábránkkal azt kívánjuk szemléltetni, hogy két paraméterrel jellemzett fonetikai elem szóráskepe esetén milyen átfedések keletkezhetnek, tehát a gépnek minél több paramétert (vektort) kell a pontosításhoz figyelembe vennie.





6. ábra

Két paraméterrel (vektorral) jellemzett beszédhang tulajdonság egyéni szórása és több személy kiejtésének szórása.

A bizonytalanul eldönthető átfedési helyeket vonalkázással kiemeltük

Tekintsük át röviden, hogyan működik egy korszerű beszédfölismerő gép. Két fontos része van: az akusztikai rögzítés és jelföldolgozás, valamint a nyelvi azonosítás és tárolás. Az első rész klasszikus hangfölvétellel és analóg-digitális (A/D) átalakítással kezdődik. Sajnos, a letapogatási szám alacsonyan tartása miatt általában 5 kHz fölötti szűrést alkalmaznak, ami egyes mássalhangzók jellegét csonkítja.

A számítógépek ezután a beérkező akusztikai jelek tulajdonságait bizonyos vektorokkal vagy Fourier-együtthatókkal jellemzik, amelyek még mindig folytonos sort alkotnak. Ezeket dolgozza föl a szerkezet Markov-moddal, vagy akár idegi hálózattal, megfelelő ötletes algoritmusok alkalmazásával bizonyos kvantált jelsorrá. Az így nyert részegységek transzformálhatók át nyelvi elemekké egy következő földolgozó egységben, immáron jóval bonyolultabb, többszörösen visszaható, erősítő és gátló elemekkel módosított és elektronikus morzsákat bőven alkalmazó osztályozó folyamatokkal. Ezután ismét folyamatos hangfűzérre alakítható a jelsor és tárolóban rögzíthető.

A gép fölismerő tevékenysége abban áll, hogy a rögzített hangcsoportok sokszori bemondás alapján középértékelt és szórásában ismert szótári egységeit új bemondással egyezteteti. Ez valószínűségi elvek, többváltozós rekurziós egyenletek, vagy egyéb összehasonlító eljárások (legkisebb eltérés) segítségével valósul meg. Tulajdonképpen ilyenkor a gép ugyanazt a munkafolyamatot végzi, amit a betanításkor minden esetben végigjárt. A tökéletesítéshez természetesen az idővetemítés és a dinamikaszugorítás is hozzátartozik. Kevés számú szótári anyag esetén ez a fajta személytől függő fölismerés jó (94-96%-os) eredménnyel jár.

Minden gép csak annyi anyagot képes fölismerni, amennyire "betanították". Az egyetlen személy kiejtésére működő fölismerő szerkezetek betanítása könnyebb és gyorsabb, de akkor is hosszadalmas, többszöri bemondással lehet az adattárat véglegesíteni. Az ún. személytől független gépek betanítása azonban nagyon körülményes. Egy francia szerző (Gauvin 1986) szótagok fölismerésére készített gépet. Rendszerében 2498 szótag és ezekből összeállított 10 400 szó szerepel. A gépen egyetlen "begyakorolt" bemondó a teljes szótaglistát és a 10 400- szavas szókészletet 2 hónapi, összesen 19 alkalommal végzett "betanítási" idő (összesen 4+13 óra) alatt rögzítette. Ugyanezen személy ellenőrző vizsgálatkor szótagokra 12%, szavakra 6% hibát vétett a gép. Ehhez a munkához a szerző a francia szavak szótagok szerinti eloszlását is figyelembe vette. Bár nem tartozik szorosan a fölismeréshez, de bemutatjuk, hogy ez a statisztika hogyan tér el saját magyar eredményeinktől (Tarnóczy 1961). Egytagú: fr. 12%, m. 39%; kéttagú: fr. 43%, m. 34%; háromtagú: fr. 36%, m. 19%; négytagú: fr. 8%, m. 6%; öttagú: fr. 1%, m. 2%. Megjegyzendő, hogy az ilyen felosztás nem is helyes, világosabban képviseli a nyelvet a szótagok fonémáinak belső eloszlási statisztikája.

A rohamos fejlődés és az elképesztően nagyszámú irodalmi adat alig teszi lehetővé a helyzet objektív áttekintését. A dolgozatok általában az IEEE szervezet és az IBM cég folyóirataiban jelennek meg, közöttük a nagy létszámú ASSP (Acoustics Speech and Signal Processing) 1976 óta évenként tartott összejöveteleinek átlag 3000-lapos nagy kvart alakú kiadványaiban. Újabban további helyi (európai, japáni) konferenciákon is ontják az új eredményeket. Összefoglalásra a fejlődés jelenlegi szakaszában különben sem vállalkozhatunk. Megemlítjük az IBM 22-tagú szerzői kollektívájának (Averbuch et al. 1986) ismertetését a TANGORA-5 fölismerő gépről (Tangora az eddigi legnagyobb sebességet elért gyorsíró neve), amely 5000 irodai használatú szót ismer föl azonosidejű visszajelzéssel, bárki ejtésében, aki a gépet előzőleg (20 perc alatt?) betanította. A gép a piacon a mai napig nem jelent meg!

Az azonosidejűségről később még megemlékezünk, van azonban ennek a gépnek egy jó tulajdonsága, éspedig az, hogy egy-, két- és háromelemes hangstatisztika figyelembevételével készült, és hogy egy kiejtési szótár is tartozik hozzá, ami az angolban sok gondot old meg.



Más kutatóhelyek, így a Haskins Laboratórium, ahol a kitűnő információelméleti matematikus, L. Rabiner is dolgozik, arra utalnak, hogy a betanítási idő a megoldandó feladat függvénye, így három napig is eltarthat. Végül is csodáltnivaló elektronikus fogások és a nagy matematikai és félvezetőtechnikai fölkészültség sok értékes és hasznosítható eredményei mellett a beszédfölismerés gépi megoldása még várat magára.

Egy 1990-re vonatkozó áttekintés szerint (Mosotto 1991) az amerikai piac 1,8 milliárd dollár termelést ért el az összes jelföldolgozó gépek forgalmazása terén. Az európai ugyanakkor csak 146 millió dollár volt. Európában a piac kétharmadát Anglia és Németország termeli, Franciaország feltörőben van. A felhasználás fő területe a szolgáltatás és az irodai alkalmazás (csaknem 75%, és növekszik), és mintegy 20%-a ma még a fölismerés, de ez csökkenni fog. A technikai megoldások a beszédfeldolgozás terén elsősorban a távközlésben alkalmazhatók, azonban lemaradás van a szöveget beszéddé átalakító, a beszédfölismerő és a személyazonosító rendszerek fejlesztésében. Ez a kritika az Eurospeech 1991. évi genovai konferenciának bevezetőjéből való. Hozzátehetjük, hogy a fő okot éppen abban látjuk, hogy még nem találták meg a folyamatos beszédjelsor kvantált elemi egységekre való alakításának átütő erejű megoldását.

Lássuk csak, hogyan foglalta ezt össze James Flanagan, a beszéd kutatás egyik jelentős személyisége Genovában (1991). Röviden azt mondta, hogy normális, párhuzamosan kapcsolt számítógépekkel, folyamatos beszéd fölismerésére a mainál 1000-szer gyorsabb folyamatokra volna szükség, vagy ha tetszik azonosidejű eredményt másodpercenként 100 milliárd vizsgálati pontú operációval érhetnénk el. Ez nagyjából annyit jelent, hogy mire mai eszközeinkkel egy mondatot valóban azonosidejű alakjában meg tudnánk a géppel értetni, maga a számítógép generáció avulna el. Pillanatnyilag az 1  $\mu$ -os technológiával a feldolgozás üteme  $10^7$ - $10^8$  művelet/s, vagyis a 15 kbit/s képességű gépek talán az 100 szavas megoldásig elvezetnek (bár ilyen még sehol sem kapható), de az ezredév végére eljuthatunk a 0,35  $\mu$ -os morzsa technológiához, mivel akár 8 kbit/s képességű géppel is elérhető lenne egymilliárd művelet/s földolgozás, vagyis 2000-ig el lehetne jutni az 5000 szavas folyamatos szófölismerésig. Mindez azonban erősen függ a morzsák teljesítménynövelésének haladásától és a fölmerülő újabb nehézségek (azonosidejűség) legyőzésétől.

#### **IV. A választott fonetikai rendszer**

Az emberi beszéd nagy terjengősségű jelsor. Több szempontból is. Ezúttal csak a fonémák terjengősségét említjük, hiszen a digitális jelrendszer (akár a Morze-abc is) mutatja, hogy a fogalmak jellemzésére sokkal kevesebb elemmel is el lehet jutni. Az ilyen rendszer azonban csak nagyon hosszú képletsorral fejezheti ki az egyszerű fogalmakat. Viszont a mai nyelvekben meglévő hangzó-



készlet (és a fonémakészlet is) kimutathatóan túl nagyszámú a feladat tisztességes megoldásához. Korábban említett adatok (Tarnóczy 1992) szerint a mai magyar fonetikai rendszerben 9464 független és jól érthető CVC kapcsolat létesíthető, de a szótári anyag csak 738 közkeletű szavat tart számon; ez a lehetőségek 8%-a! Egy redukált számú fonetikai rendszer rendkívüli mértékben egyszerűsíthetné a gépi fölismerés eddigi nehézségeit. Megkockáztathatjuk azt a kijelentést is, hogy az azonosidejű mondatfölismerés a programlehetőségek kellő átalakításával akár ma is megoldható lenne.

Alapvető kérdésnek tekintjük, hogy elemi fölismerési adatnak mit tekintünk. Az irodalomban a fonémának és a szótagnak egyaránt fontos szerepe van a fölismerési modellben. Fonéma jóval kevesebb van, mint szótag. Az ökonómia azonban emberi megértés szempontjából nem jelentős, hiszen úgyis óriási a nyelv terjedőssége (redundanciája). A szótagok fölismerésének fontosságát igazolják az érthetőségi kísérletek 100-200 ms terjedelmű fölismerési egységei. Végül megemlíthető egy tanulságos kísérlet, amelyet Colin Cherry laboratóriumunkban demonstrált a 60-as években. Halkan a fülébe (a jobb fülébe!) mondott folyamatos szöveget szótagkészsényi követéssel folyamatosan utána mondta. Magyarul, amely nyelvet nem ismerte, még hozzá elég jó kiejtéssel. Mindezek ellenére az egyszerűbb megoldást, a fonémákból összeállított rendszert választjuk, mert egyes mássalhangzók gyors lefutási ideje pszichoakusztikai kísérletek szerint folyamatos beszédben nehezen különíthető el és rögzíthető a pontos időbeli rendbe, ha nem magára az emberi agyra bízunk a feladatot. Márpedig rendszerünket beszédfölismerő gépeknek is pontosan és azonosidejű rendben kell rögzíteniük.

Az elemek kiválasztásában több szempontot kell figyelembe vennünk. Az emberiség mintegy 8000 nyelvet használ, ebből 400-nak van írott formája (Cherry 1966). Az utóbbiak közül talán a felének van jelentős irodalma, s végül mintegy 16-ra csökken a mértékadó tudományos irodalomban figyelembe veendő nyelvek száma. Az emberiség közkinccse mindez csak akkor lesz, ha 1. mindent minden nyelvre lefordítunk, 2. ha mindenki megtanul minden nyelvet, 3. ha találunk egy olyan nyelvet, amelyre minden másikat lefordítunk.

A művelt világnak (értsd Európának) volt egy nem is rövid korszaka, amikor ez az utóbbi föltétel majdnem megvalósult; gondolunk itt a görög-latin nyelvi kultúrára. A nagy irodalmi nyelvek örökértékű műveinek egymásra fordítása már a későbbi "bábeli" Európa teljesítménye, de ma már ez sem járható. Most éljük meg a tudományos nyelv hegemonizálódási törekvését – éppen nem a legjobb irányban. Miért mondom ezt? Shakespeare nyelve gazdag, kifejező, értékes. De az angol fonetika nemcsak bonyolult, hanem a kiejtés tapasztalatai szerint pontatlan. A többé-kevésbé érthető mondatokból egyes szavakat kivágva, azok teljességgel érthetetlenek. Többek közt ez az oka az angolra alapított beszédfölismerő gépek leküzdhetetlen nehézségeinek. Egy japáni fölismerő rendszer szerkezeti (hardware) része kevesebb, mint felébe kerül egy amerikainak, a program a



megoldások (software) pedig sokkal hatékonyabbak ugyanolyan eredmények eléréséhez. Még a francia gépek is sokkal egyszerűbbek, mint az amerikaiak. Példaként említhetjük, hogy az európai piacon a francia nyelvre alapított DATAVOX 5000 különálló vagy 300 folyamatos szövegű szó fölismerési képességével 70 000 frankért megvásárolható.

Ezen a helyen alkalmunk van arra, hogy a természet egy fontos alaptételére utaljunk. Ez pedig úgy szól, hogy – minden látszat ellenére – a biológiai képződmények, így maguk az élőlények is eredményesebben alkalmazkodnak a technikai kényszerhez, mint viszont. Nemcsak a hosszú távú alkalmazkodásra (az evolúcióra) igaz ez, hanem a mai kényszerhelyzetek megoldásaiban is sok példa van rá: a mikroszervezetek gyógyszerellenálló képességének kifejlődése, vagy az agrártechnikában a szárazságot tűrő populációk kitenyésztése stb. Egyik kézzelfogható mai biológiai siker, hogy miután a paradicsomszedő gépeket semmiképpen nem sikerült mindenfajta termés szedésére hatásossá tenni, kitenyésztették az egyforma nagyságú és alakú paradicsomtermést, amit egyetlen géptípus hibátlanul learat. Ezekre az analógiákra hivatkozva érdemes elgondolkozni, hogy a sokféle nyelvre kifejlesztett sokféle fölismerő gép helyett nem volna-e célszerű az egész emberiség számára egy – egyszerű gép által is könnyen megfejthető – közös nyelvet kialakítani.

A világnyelvek rossz eredményei és a japáni közlemények sikeres beszámolóí irányítják rá figyelmünket a kevés mássalhangzótorlódást tartalmazó, és szótagjaira könnyen szétválasztható nyelvek előnyös helyzetére gépi fölismerésben. Nem akarom, persze, azt mondani, hogy angol helyett a japáni nyelv legyen a tudományos világnyelv. De okvetlenül gondolnunk kell az egyszerűbb fonetikájú nyelvi rendszerek tanulmányozására. Mint az általános nyelvészetben járatlan idegen, annyit azért megemlíthetek, hogy a Csendes-óceáni nyelvek három-öt magánhangzóra épített nyelveit érdemes ebből a szempontból tanulmányozni. A 3-magánhangzós nyelvek általában 5-6 mássalhangzóval rendelkeznek, 5-magánhangzósakra pedig 7-8 mássalhangzó jut. Természetesen fölmerül az a gondolat, hogy vannak nemzetközi nyelvi próbálkozások, s közöttük a legjelentősebb az eszperantó és az ido. 1901 és 1907 között a "Délégation pour l'adaptation d'une langue auxiliaire internationale", amelynek Eötvös Lóránd is tagja volt, alaposan megvizsgálta mindkét nemzetközi nyelv alapjait, architektúráját és nyelvtanát. Jelöléstechnikailag az ido jobb (majdnem fonetikus a helyesírása), amellet a jól ismert nemzetközi szavak eredeti vagy közel eredeti formájukban megmaradtak benne. Az eszperantó nyelvtanilag alaposabban ki van dolgozva, de csak a latin nyelvrokonságot veszi figyelembe. Felmerül a kérdés: miért nem ezekre esik a választásunk? Minden tiszteletünk a csodálatos alapgondolaté és az elért eredményeké. Nemzetközi nyelvre mindinkább szükség van. Csakhogy most az eddigi szempontokhoz egy nagyon fontos újabb csatlakozik.

Akkoriban még nem volt aktuális a gépi fölismerésre való alkalmazhatóság, ezért ezek a nyelvek nem fonetikai, hanem grammatikai alapon készültek. Még így is jobbak gépi fölismerés céljára, mint akár a német, akár angol, de azért a japáni előnyöket nem érik el. Gondolhatnánk esetleg az európai élő nyelvek közül az olaszra, vagy a finnre, de a beszédhangok megkívánt könnyű szétválasztási lehetőségét ezek is csak részben tartalmazzák. Legfontosabb azonban az, hogy egyetlen élő nyelv szótára sem tartalmazza a lehetséges hang, ill. fonéma kombinációknak még a 2-3%-át sem. Tehát kevesebb elemmel és jól megkülönböztethető formákkal helyesebbnek látszik dolgozni. A gépi fölismerésben nem szerepel a redundancia; helyette az információ pontos megismerése, vagyis a beszédhang-elemek és a fonémák kölcsönös és egyértelmű megfeleltetése elegendő a teljes fölismeréshez. Magyarán: ha a gép azonosidejű formában leírja az elhangzott szöveget, megtette kötelességét.

Az új nemzetközi nyelv szükségességét éppen azért vettem föl először 1965-ben és azóta többször is, mert ezáltal nemcsak nyelvi, hanem információelméleti szempontok szerint is nagy lépést tehetünk a gyakorlati alkalmazások irányába.

Az alábbiakban egy ilyen fonetikai rendszert próbálok fölvezetni. Először, ami a magánhangzókat illeti, nem kell külön indokolni, hogy az ötelemű rendszerünk miért lesz

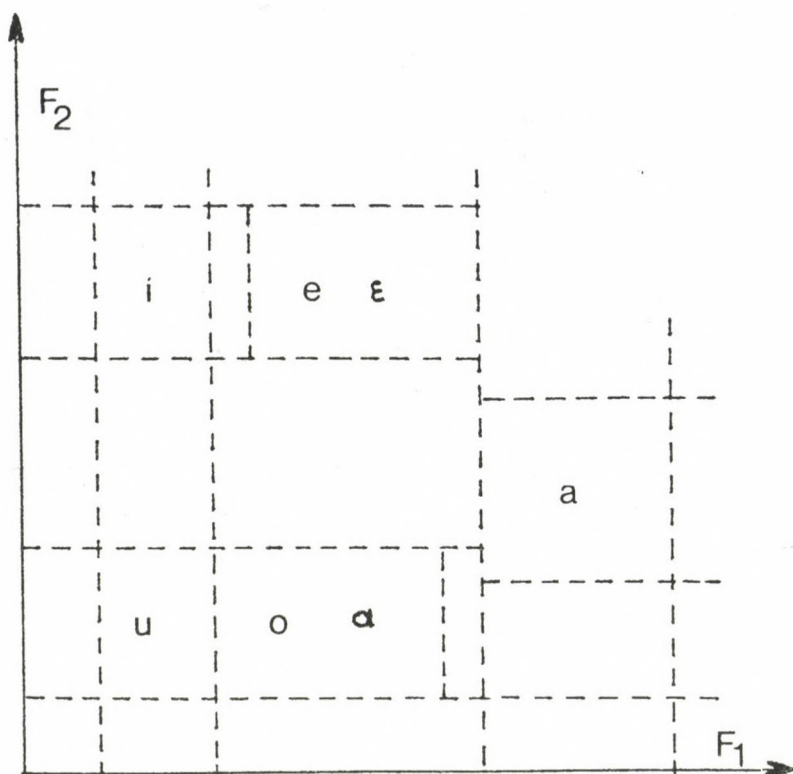
**u – o – a – ε – i**

Ha az  $F_1$ ,  $F_2$  elrendezésüket a 7. ábrán megtekintjük, azonnal látjuk, hogy a sarokmagánhangzók nagyon jól szétválaszthatók, a két közbensőnek azonban bizonytalan megoldásai is vannak. Ezért, valamint főként azért, mert célunk szerint bármilyen nyelv hangjainak ebbe a rendszerbe be kell férnie, a következő engedmények tehetők:

- az **o** egészen az **a** jellegig elhúzódhat, ami különösen a többfajta ehhez hasonló hangzók nyelvtől függő használatának problémáit oldja föl,
- az **ε** színeképe hasonlóképpen akár az **e**-ig terjedhet, szintén tekintettel a különféle variációs ejtésekre,
- semmi különbséget nem teszünk rövid és hosszú ejtési változatok között, ami némi színeképi változást is jelenthet, valamint nem vagyunk tekintettel a hangsúlyos és hangsúlytalan színeképi formák különbségére,
- az **u**, **a**, **i**, kategóriákat élesen definiáljuk, mert a tapasztalat szerint (talán a török **i** kivételével) ezek ejtése a kultúrnyelvekben egységes,
- a magánhangzók nazális ejtését mellőzzük.

A mássalhangzók összeállításában lényegesen körültekintőbb módon kell eljárunk. Bizonyos, nehezen megfogható megkülönböztető jelek egyes tévesztéseket eleve valószínűsítene. Ilyenek például a különféle nazálisok. Ezek egymástól való megkülönböztetését nem tekintjük feladatunknak.





7. ábra

Az 5-magánhangzós rendszer szűrőkörökkel elválasztható helyei az  $F_1$ ,  $F_2$  formáns szerinti koordinátarendszerben

Más esetben arra kell figyelni, hogy egyes nyelvekben nem tesznek különbséget két fonéma, például az /l/ és /r/ között. Ebben még az is zavarhat, hogy az /r/ uvuláris és elveoláris kiejtése nyelvenként kötött. Ebben a csoportban tehát széles értelemben és megkülönböztető jellegek kiküszöbölésével egy egységes likvidát definiálunk, amely esetenként akár pergő is lehet.

Semmi körülmények között nem használunk affrikátákat és félmagánhangzókat. Az affrikáták nemcsak egymás között téveszthetők, hanem a homorgán zár-, illetve réshangokkal is. További könnyítésként a mássalhangzók időtartamát (rövid-hosszú változat) sem vesszük tekintetbe, valamint azonosnak tekintjük a kiválasztott fonémák zöngétlen és zöngés változatait. A gépen úgyis 100 Hz alatti szűrést kell használnunk, mert személyazonosítással nem kívánunk foglalkozni. A félmagánhangzók ([w], [y]) megzavarnák a szótagok beosztását, hiszen élő

nyelvekben hol így, hol úgy értelmezhetők. Az  $\eta$  szóvégi vagy vokálisközi megjelenése sok nyelvben gyakori. A javasolt sorozat tehát:

$s - \int - p - t - k - m - l - \eta$

amivel kapcsolatban a következő rendező elvek rögzítése látszik szükségesnek:

- a réshangok közül az /f/ részben kis intenzitása, részben jellegtelen (fehérzaj jellegű) színe miatt maradt ki,
- a három zárhang megkülönböztető jellege ugyan elég nehézkesen tisztázható, de ezt a mai fölismerő szerkezetek is megoldják, viszont fontos érv mellettük, hogy a folyamatos hangfűzér gépi tagolásában (szegmentálásában) főszerepük van,
- az **m** nazális jelleget képviseli, bármikor **n**-nel helyettesíthető,
- az **l** – **r** azonossági jellegét említettük,
- az  $\eta$  magánhangzó után magában álló mássalhangzó.

Az összeállított szavak kiejtésének egyszerűsítésére, valamint a gépi fölismerés hatékonysága érdekében a hangkapcsolatok tiltó szabályai a következők:

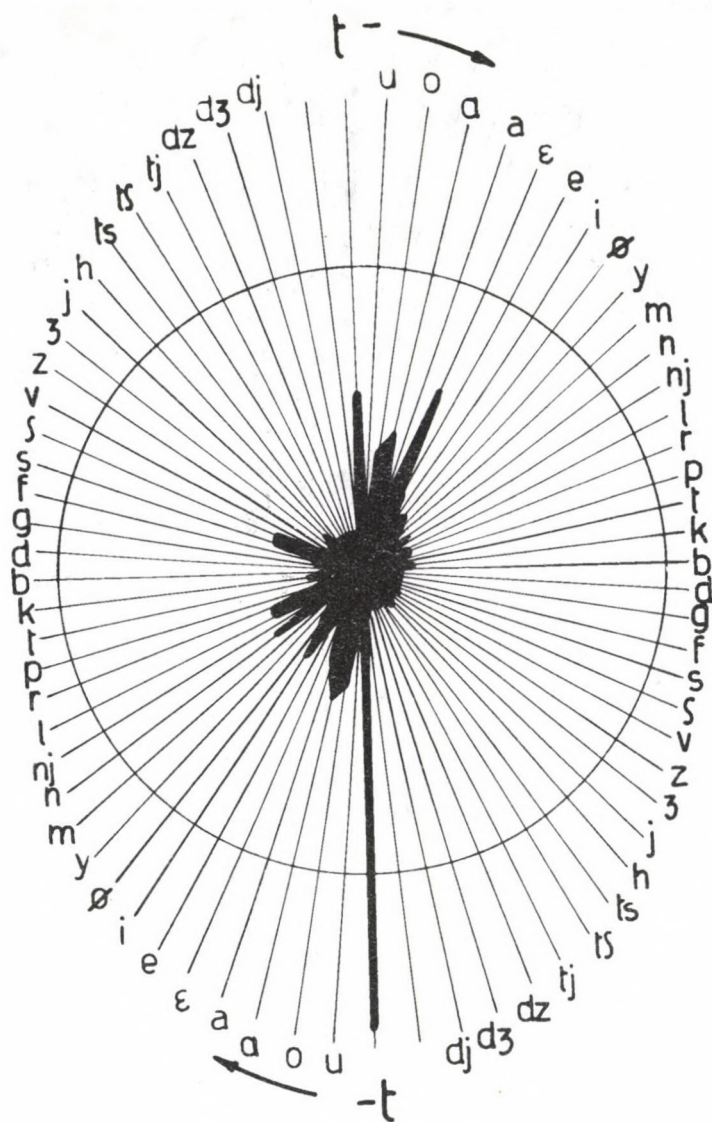
- VV kapcsolódás tilos
- CCC kapcsolódás tilos
- CC kapcsolatban nem kerülhet egymás mellé két réshang,
- CC kapcsolatban tilos a **t** után bármelyik réshang
- CCV első hangja nem lehet **m** és **l** és **t**
- CCV az első **p**, és **k** hang után csak  $\int$  és **l** állhat
- VCC utolsó hangja nem lehet **m** és **l**
- VC és VCV kapcsolat mássalhangzója  $\eta$  is lehet.

A legkényesebb kapcsolódások a **p**, **t**, **k** körül alakulhatnak ki. Ezért a 8.-10. ábrán a magyar nyelvre összeállított statisztikánk szerinti kapcsolódási arányokat szemléltetjük. Az ábrákon nagyobb magyar nyelvű szöveganyagból statisztikázott kapcsolódások gyakoriságokat tüntettük föl (fekete tuskék). Az arányokhoz figyelembe kell vennünk, hogy maguk a vizsgált hangok (valóságban betűk) is bizonyos gyakorisággal fordulnak elő a nyelvben. Néhány következtetés így is levonható. Magyar nyelvre ez a képsorozat minden mássalhangzóra megvan, más nyelvekre is érdekes volna elkészíteni.

Megjegyezzük, hogy mostani összeállításunk fonetikus jelölésnek tekinthető, ezért hagytuk ki például az affrikáta-gyanús kapcsolatokat. Valószínűleg arra is gondolni kell még, hogy mondatok összeállításakor a szavak közötti mássalhangzótorlódások ne okozzanak zavart, ezért a VCC végződő és CCV kezdő jellegű szavak lehetőleg ne kerüljenek egymás utáni rendbe.







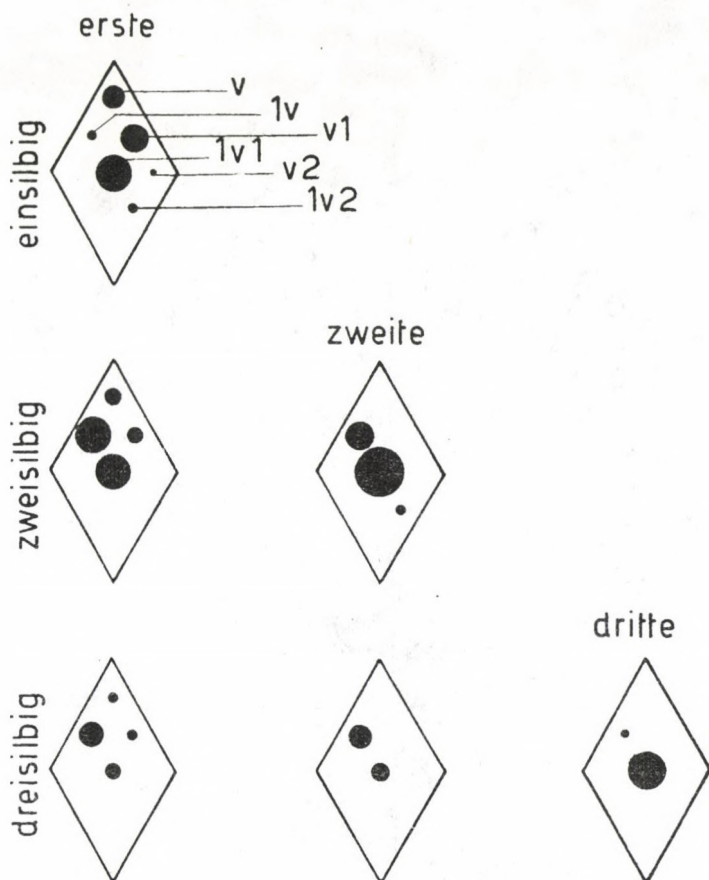
9. ábra

A **t** hang kételemű kapcsolatai a magyar nyelvben.

A **p**-vel szemben nagy (7,7%) a gyakorisága, főként végső hangként fordul elő, és majdnem minden mássalhangzóval gyakran kapcsolódik (múlt idő)







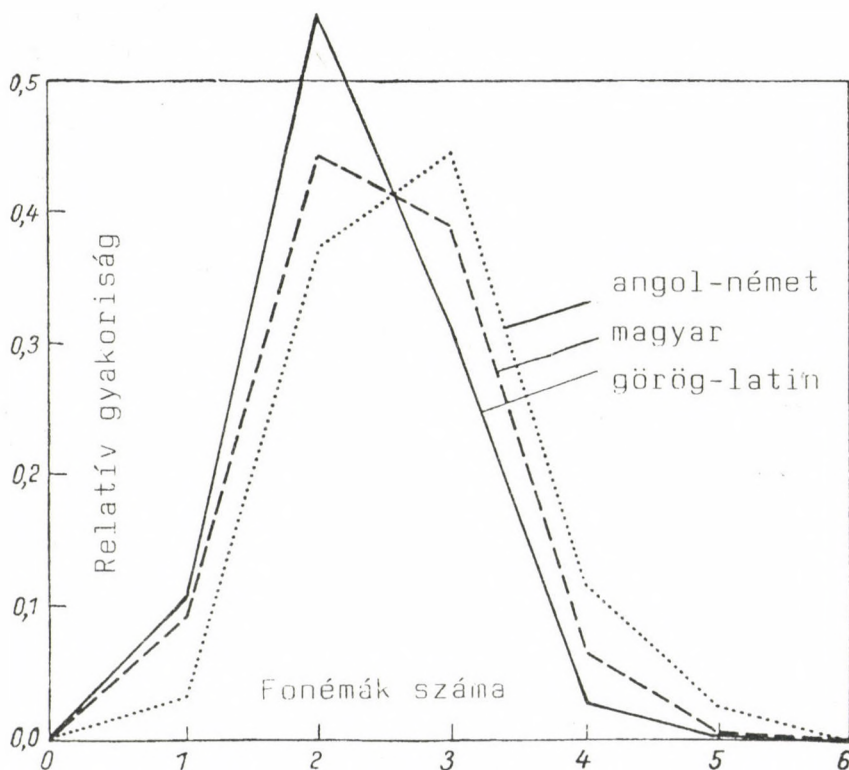
11. ábra

Egy-, két- és három szótagú szavak szótageloszlását jelképező Menzerath-féle rombuszok magyar nyelvre. A fekete körök arányosak a megfelelő szótagtípusok előfordulási gyakoriságával. Az elemek száma fölülről lefelé eggyel növekszik, középen mindig V áll

Fontos kérdés a szótagtípusok nyelvi eloszlása is. Ezekre vonatkozólag közöljük a 11. ábrát, szintén magyar nyelvi statisztika alapján (Tarnóczy 1961). Az ábrából láthatólag a leggyakoribb CVC alakú szótagok mellett (az utolsó szótag 76%-a ilyen), a kétszótagú szavak első szótagja gyakran CV alakú, és teljesen hiányzik a CCV. Más nyelvekre a szótagtípusok más alakúak, például ez a sajátosan magyar arány a németben a CCV javára erősen megváltozik. Véleményünk szerint mégis minden fajta megengedett kapcsolat létesítése indokolt, nagy eltérések az egyes nyelvek között nem valószínűsíthetők. A szótagolás sem minden nyelvben azonos, hiszen négyelemű szótagok is előfordulnak, különösen az an-



golban és a németben (12. ábra, Tarnóczy 1960). Végül a 12 elemű rendszerünk két- és háromtagú kapcsolati alakzatait, valamint a lehetséges maximális szókészletét mutatjuk be a 3. táblázatban. Kéttagú, magánhangzóval kezdődő szó 23 ezer, háromtagú 5 millió készíthető, a mássalhangzóval kezdődő kéttagúakból fél millió, a háromtagúakból pedig 120 millió is összeállítható. Ötmagánhangzós rendszerünk eszerint túldimenzionált az elérendő célhoz képest. Nagy valószínűség szerint egy 3 magánhangzós és 5 mássalhangzós rendszer is megtenné, nem beszélve arról, hogy a gépi fölismerés abban már gyerekjáték volna. Viszont igaz, hogy hosszabb szavak alakulnának ki.



12. ábra

A szótagokat képező hangok (fonémák) számának eloszlása három nyelvcsoporthban. A magyar saját eredmény 1960-ból, a másik kettő W. Fuchs 1956-os adataiból összevonva. Vízszintes tengely a szótagra eső hangok száma, a függőleges tengely nagyjából a gyakoriság törtrésze

3. táblázat

Kéttagú V-kezdetű		
VCV	200	Ali
VCVC	1 600	eper
VCVCC	2 200	áraszt
VCCV	950	irtó
VCCVC	7 600	életet
VCCVCC	10 450	osztást
	23 000	

Háromtagú V-kezdetű		
VCVCV	8 000	iráni
VCVCVC	64 000	utasok
VCVCVCC	88 000	italunk
VCCVCV	38 000	iktató
VCCVCVC	304 000	értelem
VCCVCVCC	418 000	elterelt
VCVCCV	38 000	okolta
VCVCCVC	304 000	átértem
VCVCCVCC	418 000	édenkert
VCCVCCV	180 500	elsózta
VCCVCCVC	1 444 000	alkalmas
VCCVCCVCC	1 985 000	ártatlant
	5 310 000	

Kéttagú C-kezdetű		
CVCV	1 400	kupa
CVCVC	11 200	korong
CVCVCC	15 400	bevált
CVCCV	6 650	parti
CVCCVC	53 200	portás
CVCCVCC	93 150	felfest
CCVCV	3 200	pláne
CCVCVC	25 600	kráter
CCVCVCC	35 200	stoppolt
CCVCCV	15 200	Spárta
CCVCCVC	121 600	flörtöl
CCVCCVCC	162 200	krampuszt
	514 000	

Háromtagú CV-kezdetű		
(a V-kezdetűek 7-szerese)		

Háromtagú CCV-kezdetű		
(mintegy 85 millió)		

Lehetséges hangkapcsolatok száma az ötmagánhangzós rendszerben  
(magyar példaszavakkal szemléltetve)

Természetesen új műnyelv nyelvi rendszerét is erősen át kell gondolni, például, hogy ragozó, hajlító vagy két- és háromtagú állandó tövű formája legyen-e, tartsuk-e meg a latin mintájú nemeket stb.? Ugyanakkor könnyen meg lehessen tanulni, s ebben ne annyira az eszperantóhoz, mint inkább a különböző nyelvekből jól összeválogatott szókészlettel induló ido rendszerhez hasonlítson-e? Olyan szabályok is segítenének az elsajátításban és a gépi fölismerésben, hogy pl. az egyszótagnak (táblázatban nem szerepel, de 2000 ilyen szó képezhető) kötőszók, névelők, gyakori határozószók legyenek, vagy hogy a főnevek V-vel, az igék C-vel kezdődjenek, vagy hogy az alapszavak kéttagú, a származtatott vagy képzett szavak háromtagú formában jelenjenek meg, vagy bármely más információtechnikai fogás, ami a gépi fölismerést könnyíti. Fontos lehet a latin rendszerű három nemű egyeztetés megtartása, ami a mellékmondati visszautalásokat nagyban



megkönnyíti. Mindez alapos tervezési munkát igényel, ami elsősorban számítástechnikában járatos, általános nyelvészetben és fonetikában kitűnő ismeretekkel rendelkező nyelvtudósoknak, vagy ilyen személyek munkacsoportjának lehet a feladata. Az azonban biztos, hogy a számítógépes beszéd-fölismerést megoldaná.

Még meg szeretnék említeni egy időben előremutató gondolatot. Ha ez a megfelelően kidolgozott, és minden gondolati finomságot visszaadni tudó, amellett kiejtésben könnyű és számítógép által is érthető nyelv gyakorlatilag beválna, előbb-utóbb mindenki megtanulná. Akkor pedig mindenki számára elég volna a saját és az univerzális világnyelv tudása; a fordítógépeknek is csak ezt a két fordítási feladatot kellene ellátniuk, és a világ teljes irodalmát egyetlen nyelven lehetne élvezni. És ez sem kevés!

### Köszönetnyilvánítás

Köszönettel tartozom Ádám György professzornak, aki nemcsak átnézte, hanem tanácsaival lényegesen módosította az agytevékenység ismertetésével foglalkozó részt. Ugyancsak köszönöm Gósy Mária kandidátusnak a nyelvi jelölések és kifejezések pontosítása terén adott tanácsait. És köszönöm a kézirat lassú kialakításával kapcsolatos késés iránti kiadói türelmet.

### Irodalom

Csak válogatott anyag közlésére van lehetőség, ezért a II. részben néhány klasszikus értékű fíziológiai könyvön kívül (ahol lapszámokra nem is hivatkozom), elsősorban a dolgozatban említett fontosabb klasszikus eredmények és áttekintések felsorolására szorítkozom. A III. részhez az 1976 óta évente megjelenő ICASSP négykötetes előadásanyagában mintegy 8000 idevágó cikk található.

Ádám György: Érzékelés, tudat, emlékezés. Budapest 1976, 192–210.

Averbuch, A. et al. (21 Co-authors): An IBM PC Based Large-Vocabulary Isolated-Utterance Speech Recognizer. ICASSP 86 Tokyo, 1986, 53–56.

Békésy, G. v.: Über die Resonanzkurve und die Abklingzeit der verschiedenen Stellen der Streckentrennwand. Akust. Z. 8, 1943, 66–76.

Blauert, J.: Räumliches Hören. Stuttgart 1974, 60–61.

Cherry, C.: On Human Communication. 2nd ed. Cambridge Mass. 1966.

Cerebral Cortex Vol. 3. Ed: E. G. Jones and A. Peters. New York – London 1985.

Cochlear Implants. Ed: R. A. Schindler and M. M. Merzenich. New York, 1985.

Feddersen, W. E. et al.: Localization of high-frequency tones. JASA 29. 1957, 988–991.

Flanagan, J. L.: Speech Analysis, Synthesis and Perception. 2nd ed. Berlin – Heidelberg – New York 1972.

Flanagan, J. L.: Speech Technology and Computing. A Unique Partnership. Opening Session, Eurospeech. Genova 1991, 7–22.

Fletcher, H. Speech and Hearing in Communication. Van Nostrand, New York 1953, 415–423.

Fletcher, H. – Steinberg, J. C.: Articulation testing methods. Bell Syst. Tech. J. 8. 1929, 806–854.

Forssmann, W. G. – Heym, Chr.: Neuratanomie. 3-te Aufl. Berlin – Heidelberg – New York 1982.

- Fucks, W.: Die mathematischen Gesetze der Bildung von Sprachelementen aus ihren Bestandteilen. *Nachrichtentech. Fachber.* **3.** 1956, 7–21.
- Gábor, D.: *Theory of Communication.* J. Inst. Elec. Eng. **93.** III. London 1946, 429.
- Gauvin, I.-L.: A Syllable Based Isolated Word Recognition Experiment. ICASSP 86 Tokyo, 1986, 57–60.
- Kimura, D.: Some effects of temporal damage on auditory perception. *Canada. J. Psychol.* **15.** 1961, 156–165.
- Kimuran, D.: Left-right differences in the perception of melodies. *Quart. J. Exp. Psychol.* **16.** 1964, 355–358.
- Kimura, D.: Functional asymmetry of the brain in dichotic listening. *Cortex* **3.** 1967, 163–178.
- Lenneberg, E. M.: Understanding language without ability to speech: A case-report. *J. Abnorm. Soc. Psychol.* **65.** 1962, 419–425.
- Licklider, J. C. R. – Pollack, I.: Effects of differentiation, integration and infinite peak clipping upon the intelligibility of speech. *JASA* **20.** 1948, 42–51.
- Licklider, J. C. R. – Miller, G. A.: The perception of speech. In: S.S. Stevens. *Handbook of Experimental Psychology.* New York 1951, 1040–1074.
- Miller G. A. – Licklider J. C. R.: The intelligibility of interrupted speech. *JASA* **22.** 1950, 167–173.
- Miller G. A. – Nicely, P. E.: Analysis of perceptual confusions among some English consonants. *JASA* **27.** 1955, 338–342.
- Miller, J. D. et al.: Discrimination and labelling of noise-burst sequences with varying noise-laid times: An example of catagoral perception. *JASA* **60.** 1976, 410–417.
- Moore, B. C. J.: *An introduction of the psychology of hearing.* London 1982.
- Mosotto, C.: *Speech Technology and Telecommunication.* Opening Session, eurospeech. Genova 1991, 23–39.
- Ojemann, G. A.: The intrahemispheric organization of human language, derived with electrical stimulation technique. *Trends Neurosci.* **6.** 1983, 184–189.
- Pickles, J. O.: *Physiology of Hearing.* London 1982.
- Pujol R.: *Neurobiology of the cochlea.* Colloque de Physique Vol. I. Lyon 1990, 99–105.
- Rabinowitz, W. M. et al.: Relations among different measures of speech reception in subjects using cochlear implants. *JASA* **92.** 1992, 1869–1881.
- Ribári, O.: Hearing improvement achieved by different cochlear implantation techniques on the same and contralateral ears. *Acta Otolaryng* **112,** 1992, 225–229.
- Schmidt, R. F. et al: *Grundriß der Siennesphysiologie.* 4-te Aufl. Berlin – Heidelberg – New York 1980.
- Seldon, H. L.: The Anatomy of Speech Perception. Human Auditory Cortex. In: E. G. Jones – A. Peters ed.: *Cerebral Cortex.* Vol. 3. New York – London 1985, 273–327.
- Stevens, K. N. – House, A. S.: *Speech Perception.* In.: J. V. Tobias, *Foundation of Modern Auditory Theory.* New York 1972.
- Stevens, K. N.: Acoustic correlates of some phonetic categories. *JASA* **68.** 1980, 836–842.
- Tarnóczy Tamás: Az érthetőségvizsgálatok magyar szövegmintái. *Híradástechnika* **5.** 1954, 43–51.
- Tarnóczy, T.: Phonetische Gesichtspunkte bei Zusammenstellung von Texten für Verständlichkeitsmessungen. *Zs. Phonetik* **14.** 1961, 74–87.
- Tarnóczy, T.: A method for constructing phonetically balanced word list to be used for intelligibility measurements. *Symp. Intellig. de Parole, Liège* 1973, 185–190.



- Tarnóczy, T.: Über die sprachbegrenzenden Eigenschaften der Signalverteilung und der Informationsinhalte. *Zs. f. Phonetik* **13**. 1960, 57–75.
- Tarnóczy, T.: Verständlichkeits-Stabilität, Konfusiosmatrix und Verfehlsindex. *Zs. f. Phonetik* **17**. 1964, 333–338.
- Tarnóczy, T.: Can the Problem of Automatic Speech Recognition be Solved by Analysis Alone? Invited Paper. 5 ICA-Congress Vol. II. Liège 1965, 371–387.
- Tarnóczy, T.: The Formation, Analysis and Perception of Hungarian Affricates. In: Shannon and Sockey ed.: *Honor of Ilse Lehiste*. Dordrecht 1987, 255–270.
- Tarnóczy Tamás: Az emberi üzenetváltásról. *Linguistica Series C Relations* 5. Budapest 1992.
- Warren, R. M.: Perceptual restoration of missing speech sounds. *Science (N.Y.)* **167**. 1970, 392–393.
- Warren, R. M.: Auditory temporal discrimination by trained listeners. *Cogn. Psychol.* **6**. 1974, 237–256.

# HANGIDŐTARTAMOK SZÁMÍTÓGÉPES ELEMZÉSE A BESZÉD RITMIKAI SZERKEZETÉNEK VIZSGÁLATÁHOZ

Olaszy Gábor

MTA Nyelvtudományi Intézete, Budapest

## Bevezetés

A beszéd időszerkezetének vizsgálata során – annak ellenére, hogy a beszéd szerkezetének vizsgálatában ez az egyik legkutatottabb terület – még ma is sok olyan kérdés merül fel, amire nincs egyértelmű válasz. Önálló kötetben legutóbb Kassai foglalta össze e kutatási terület legfontosabb munkáit, valamint saját mérési eredményeit a magyar beszédhangok időtartamviszonyairól (Kassai 1979). A magyar beszédhangok belső időszerkezetét számítógépes szintézissel kontrollált vizsgálatok segítségével a közelmúltban vizsgáltuk (Olaszy 1985; 1989).

A beszéd folyamatjellegét figyelembe vevő időszerkezeti vizsgálatok szükségyszerűségét többek között a mesterséges beszédelőállítás megvalósulása hozta felszínre. Az analízisből elvégzett újraszintézis megmutatta, hogy az időszerkezet helyes beállítása és állandó újbóli korrigálása a beszédelőállítás egyik legkritikusabb része. René Collier így ír 1992-ben erről a problémáról: **"Manapság sok text-to-speech rendszer jól érthető beszédet produkál mondatszinten, de ebből a beszédből hiányzik a természetesség, az emberi hangszínezet, a kifejezőképesség. Nem valószínű, hogy egy hétköznapi hallgató képes lenne hosszú szövegeket úgy végighallgatni a gép felolvasásában, hogy ne lenne megerőltető számára a megértés"** (Collier et al. 1992, 32). A tanulmányban kifejti, hogy sok kutatásra van még szükség, főleg az időszerkezet területén, hogy a gépek folyamatos beszédét az emberi ejtéshez közelállóvá tehessek. Olyan szabályrendszer kialakítása tehát a cél, amelynek alkalmazásával a természetes ejtést megközelítő ritmikai szerkezetet lehet gépi úton is előállítani.

A magyar beszéd vonatkozásában folytattunk már ilyen vizsgálatokat. Az elmúlt évtized alatt kifejlesztett MULTIVOX beszédelőállító rendszerben (Olaszy et al. 1992) már működik néhány általános érvényű, időszerkezetet módosító szabály.

**Hangrövidülési szabályt** alkalmaztunk a névelőkre, ami azt jelenti, hogy a névelő magánhangzójának időtartamát rövidebbre állítottuk be, mint az *a* hangét a szöveg más szavaiban.

**Hangnyújtási szabályt** alkalmaztunk például az egyszótagú kérdések magánhangzójára. Mérések és meghallgatásos vizsgálatok egyaránt igazolták, hogy ezeknek a magánhangzóknak az időtartama lényegesen hosszabb, mint az ugyan-



ilyen környezetben lévő kijelentő intonációjú hangsorokban lévőké (Olaszy 1989). Így az ilyen kérdésekben kb. 50%-kal megnyújtottuk a magánhangzó időtartamát (pl. *Én? Most? Ő?* ).

**Összekapcsolási szabály** határozta meg, hogy a névelőket, a kérdő- és tagadószókat és az őket követő szót (egyes esetekben az előttük lévőket) egy összefüggő szekvenciának kell tekinteni és a két szót a folyamatos artikulációnak megfelelő előállításban (szünet nélkül a két szó között) kell a szintézis során kimondatni (pl. *azasztal; amadár; Mikorindulunk?; Hogyvagy?; nemtudom; nemmegyek; senkisé*m).

Az ilyen típusú szabályok hatására ritmikailag határozottan javult a hangzás, de a gépiesség még mindig erősen érezhető maradt. Ez abból is fakad, hogy a jelen rendszerben a hangsor felépítésénél mindig ugyanazokat a hangkapcsolati építőelemeket használjuk fel, függetlenül attól, hogy az adott hangkapcsolat a szóban hol szerepel, illetve hogy a szó a mondatban hol helyezkedik el. Ez azt eredményezi, hogy a **"mindig ugyanazon elemek" túl szabályos idő- és frekvenciaszerkezetet hoznak létre**, ami a hangzásban is szabályosságban jelentkezik. Az ember beszédére ilyen gépies szabályosság nem jellemző.

### A vizsgálatok célja

1. Olyan szabályszerűségeket keresünk a beszédjel időszerkezetében, amelyek nyelvi szinten mint beépült automatizmusok alakítják a hangok, hangkapcsolatok, hangsorrészek stb. időtartamviszonyait.

2. Pontos adatokkal akarjuk jellemezni az egyes kifejezési formákhoz tartozó szupraszegmentális szerkezetekben az időszerkezet egységeit, azok részelemeit.

3. Az eredményeket az automatikus beszéd-szintézisben, valamint gépi beszédfelismerési feladatoknál kívánjuk elsősorban felhasználni.

### A nyelvi anyag

A nyelvi anyagot újságból vett hírjellegű közlésekből állítottuk össze. A vizsgálatban 20 mondat adatait dolgoztuk fel egyetlen férfi adatközlő bemondásában.

A feldolgozott mondatok jellemzői:

	egy mondatban	összesen
szószám	4–8	105
szótagszám	14–27	344
hangszám	32–53	760

Az átlagos artikulációs sebesség:

14,5 hang/s

A bemondó férfi átlagos hangmagassága: 117 Hz

Néhány példa az adatbázis mondataiból:

*Pozsgay Imre lemond a rádió és televízió felügyeletéről.*

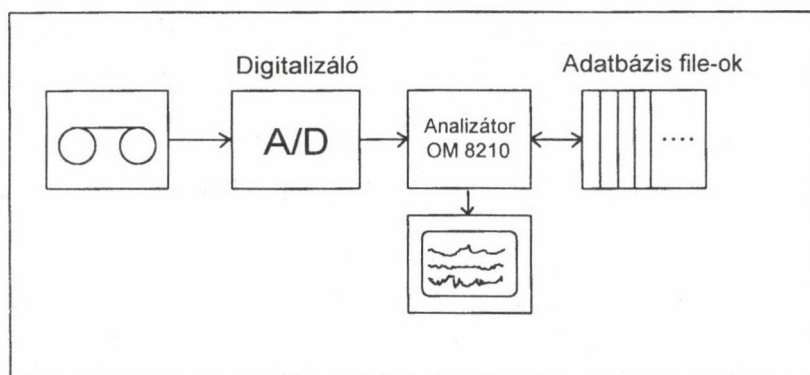
*A felügyeleti jogot a kormány bízta más kormánytagra.*

*A felügyelőbizottság teljes feloszlatására kell törekedni.*

## Módszer

### 1. Adatbázis létrehozása

Korszerű beszédakusztikai kutatás csak számítógépen tárolt, a hangzó beszédet tartalmazó adatbázisra, valamint a hozzá tartozó kezelő és feldolgozó rendszerre alapozva végezhető. Ezért jelen kutatáshoz megkezdtük egy kísérleti adatbázis létrehozását az MTA Nyelvtudományi Intézetének Fonetikai Laboratóriumában (1. ábra). Ezt az adatbázist egy későbbi, országos igényeket kiszolgáló beszéd-adatbank előkészítésének tekintjük.



1. ábra

Stúdió magnetofon, AD konverter, Philips analizátor, számítógép

Az adatbázis mondatainak feldolgozását (digitalizálás, fonetikai analízis, tömörítés, tárolás) Philips gyártmányú OM8210 típusú beszédanalizátorral és a hozzá tartozó szoftverrel végezzük. Ez a feldolgozó rendszer lehetővé teszi a beszédjel finom szerkezetű, fonetikai paraméterekkel reprezentált vizuális és auditív egyidejű vizsgálatát (Philips 1988). A feldolgozás menete a következő. A mondatokat 10 000 Hz-es mintavételezéssel digitalizáljuk, majd minden mondat beszéd-időfüggvényét 8,8 ms-os lépésekben analizáljuk, és az analízis eredményét file formájában számítógépben tároljuk. (Az alapegységnek számító 8,8 ms-os lépésegység biztosítja, hogy a beszédjel belső időszerkezetét megfelelő pontossággal írjuk le.) Ez a feldolgozási forma a beszédjelre alkalmazott egyfajta tömörítés, ami azt jelenti, hogy a vizsgálatokhoz nem az eredeti időfüggvényt használjuk, hanem annak fonetikai paraméterekkel (formánsok, sávszélességek,



energia, alaphangmagasság, gerjesztés) reprezentált változatát. Ezekből az adatokból újrasszintetizálással az eredeti időfüggvényhez igen hasonló jel állítható elő, amely az eredetileg ejtett forma összes szupraszegmentális tulajdonságát tartalmazza. Az adatbázisban egy-egy ilyen fonetikai formában tárolt mondat (file) képezi a kutatási munkapéldányt a további vizsgálatokhoz.

## 2. Az időszerkezet mérése

Az időszerkezet analízisének a számítógép támogatásával manuálisan történik. A vizsgálandó mondatot előhívjuk az adatbázisból, így az megjelenik a képernyőn összes fonetikai paraméterével (formánsok, sávzsélességek, gerjesztés, alaphangmagasság, amplitúdó) 8,8ms-os szegmenssorozat formájában. Ehhez a képernyőn megjelenő vizuális fonetikai információhoz (hasonló, mint egy hangspektrogram) kapcsolhatjuk hozzá a mindenkor auditív meghallgatást. Többféle meghallgatási forma áll a kutató rendelkezésére. Az analízáló rendszer lehetővé teszi a beszédjel szegmensein való folyamatos "görgetést" balról jobbra hangról hangra (illetve visszafelé), és a szegmensek tartalmának a görgetéssel együtt történő folyamatos meghallgatását. Ezzel a kombinált vizuális-auditív módszerrel a kutató jó hatásfokkal meg tudja állapítani a hang, szótag, szó stb. határokat és így közvetlenül időtartamokat mérhet. A munka során minden mondatban megmértük az összes hang és szó időtartamát. Az adatgyűjtés után pontos és reprodukálható adatok reprezentálják a vizsgált mondat időszerkezetét.

A rendszer biztosítja még a kapuzást is (pl. meghallgatás kijelölt ponttól kijelölt pontig), illetőleg a módosítást az akusztikai jelben (formánsok eltolása, hangok nyújtása, rövidítése, dallammenetek megváltoztatása stb.), így másfajta vizsgálatokhoz (pl. dallamgörbék, hangnyúlások stb. előállítás) is jól használható.

## Eredmények

A 20 mondat feldolgozása során nyert adatokból a következő megállapításokat tettük.

### Időszerkezet a mondat szintjén

A mondaton belül a hangfolyam folyamatos, az időfüggvényen a szóhatárok nem láthatók, vagyis az artikuláció folyamatos, a szavak egymásba folynak. A szöveg szószintű tagolására szolgáló információ elsősorban a hangmagasság-változásban található meg. Például a *Pozsgay Imre* szövegrész ejtésénél a szóvégi és szóeleji *i* hang teljesen összeolvad (időtartama megnyúlik), viszont a második szó kezdetét egyértelműen a hangmagasság növekedéséhez (ami jelen esetben az első szótagi hangsúly jelenlétét reprezentálja) lehet kötni. Hasonló jelenség mutatható ki a *rám bízta a rádió* hangsor szóvégi *a* hangja és a névelő között. Mindkét esetben mintegy 10Hz-nyi emelkedés mérhető az adott hangkapcsolat második hangjának kezdetén. A második esetben érdekes ellentmondással talál-

kozunk, mivel ez az emelkedés a névelőre esik, holott a névelőnek hangsúlytalanak kellene lennie. Ha azonban megnézzük a frekvencia-adatokat, akkor a névelőben 100 Hz-et mérünk, ami viszont megfelel a hangsúlytalan állapotnak, mert az azt követő *rádió* szó első szótagjában a hangmagasság 115 Hz-es. Ebből következik, hogy a két *a* hang közötti határjelző megkülönböztetést biztosító automatizmus úgy működik, hogy a *rám bizta* hangsor utolsó hangját még a névelő alapfrekvenciájánál is alacsonyabbra, 90 Hz-re állítja be.

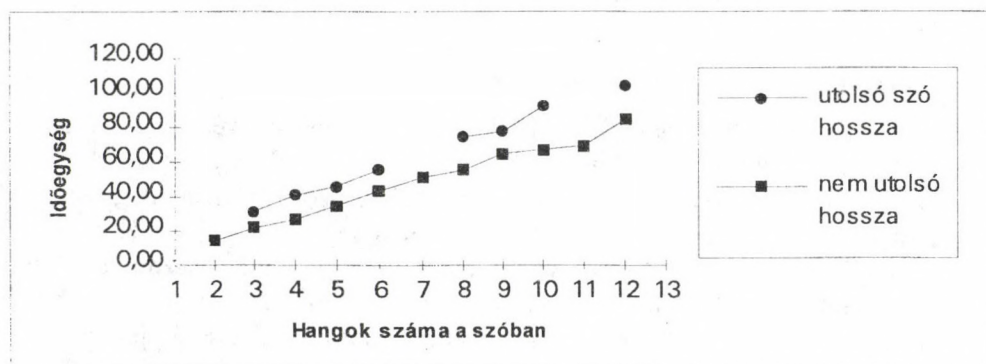
### Időszerkezet a szó szintjén

A mondatok szavainak hosszát mérve lineáris emelkedés tapasztalható a hangok számának növekedésével. Az 1. táblázatban a különböző hosszúságú szavak (hang/szó) mért fizikai hosszúságát adjuk meg. Külön kategória, ha a szó a mondat utolsó szava.

1. táblázat

a hangok száma a szóban	1	2	3	4	5	6	7	8	9	10	11	12	13
	Az időegységek száma a szóhossz kifejezésére												
utolsó szó hossza	-	-	32	41	46,5	56	-	74,6	77,6	92	-	104	-
nem utolsó szó hossza	7,6	14,8	22,3	27,2	35,3	43,8	51,6	55,6	65	67,4	69,6	84,3	-

A táblázatban szereplő számok hosszegységben fejezik ki az adott szó hosszát, vagyis például egy három hangból álló szó fizikai hossza 22,3 egység, ami  $22,3 \times 8,8$  ms azaz 196 ms. Ugyanilyen három hangból álló szó hossza, ha az a mondat utolsó szava 32 egység, vagyis  $32 \times 8,8$  ms azaz 281 ms.



2. ábra

A szó időtartamának alakulása a szóhosszúság függvényében



Az adatokból és a függvényes ábrázolásból (2. ábra) látszik, hogy a szavak relatív hosszúságát lényegesen nem befolyásolja a mondatban elfoglalt helyzetük. Egyedül az utolsó szó esetében láthatunk konzekvensen lényeges megnyúlást.

Az átlagos szóhosszúságot az alábbiakban foglaljuk össze:

1 hang/szó	(16 szó átlaga)	7,6	egység, azaz	67 ms
2 hang/szó	(9 szó átlaga)	14,8	egység, azaz	130 ms
3 hang/szó	(10 szó átlaga)	22,3	egység, azaz	196 ms
<b>utolsó szó</b>		32		<b>281 ms</b>
4 hang/szó	(5 szó átlaga)	27,2	egység, azaz	239 ms
<b>utolsó szó</b>	(2 szó átlaga)	41		<b>360 ms</b>
5 hang/szó	(6 szó átlaga)	35,3	egység, azaz	310 ms
<b>utolsó szó</b>	(2 szó átlaga)	46,5		<b>409 ms</b>
6 hang/szó	(14 szó átlaga)	43,8	egység, azaz	385 ms
<b>utolsó szó</b>		56		<b>492 ms</b>
7 hang/szó	(5 szó átlaga)	51,6	egység, azaz	454 ms
8 hang/szó	(6 szó átlaga)	55,6	egység, azaz	489 ms
<b>utolsó szó</b>	(3 szó átlaga)	74,6		<b>656 ms</b>
9 hang/szó	(9 szó átlaga)	65	egység, azaz	572 ms
<b>utolsó szó</b>	(3 szó átlaga)	77,6		<b>682 ms</b>
10 hang/szó	(7 szó átlaga)	67,4	egység, azaz	593 ms
<b>utolsó szó</b>	(2 szó átlaga)	92		<b>809 ms</b>
11 hang/szó	(5 szó átlaga)	69,6	egység, azaz	612 ms
12 hang/szó	(3 szó átlaga)	84,3	egység, azaz	741 ms
<b>utolsó szó</b>	(2 szó átlaga)	104		<b>915 ms</b>
14 hang/szó	(3 szó átlaga)	95,3	egység, azaz	838 ms
<b>utolsó szó</b>		116		<b>1020 ms</b>
15 hang/szó	(2 szó átlaga)	95,5	egység, azaz	840 ms
<b>utolsó szó</b>		106		<b>932 ms</b>

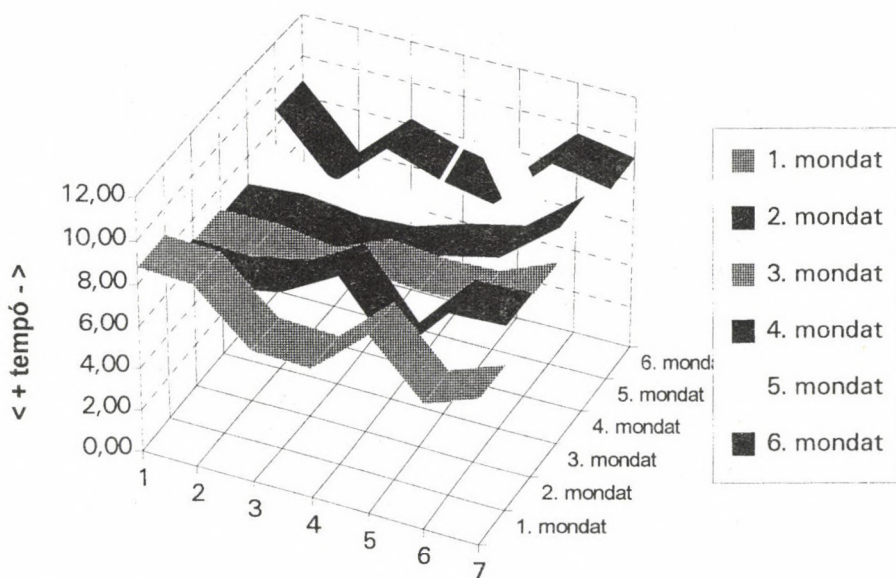
A fenti statisztikai összegzéssel egy átlag szóhosszúságot kaptunk az 1 hang/szó hosszúságú szótól egészen a 15 hangot tartalmazó szavakig, amelynek révén adott szó időtartama a szóban szereplő hangok száma alapján előre jelezhető.

A szavak kiejtési tempóját külön vizsgáltuk. Ezt a szó hangsúlyozása, valamint a mondatban elfoglalt helyzete befolyásolhatja. A vizsgálathoz hat egyforma szószámú mondatot választottunk ki, és azok **minden egyes szavában** meg-

mértük az egyes hangokra jutó elemi időegység számát (2. táblázat). A mondatok szótagszáma 21–23, a hangok száma 46–50 közötti, tehát a mondathosszúság a vizsgált mondatokban lényegében azonosnak mondható. A vizsgálat eredményét a 2. táblázatban foglaltuk össze.

2. táblázat

a vizsgált mondatok	első szó	második	harmadik	negyedik	ötödik	hatodik	hetedik
	Időegység/hang a szóban						
1. mondat	8,8	8,7	6,3	6,2	8,5	6	7
2. mondat	7,1	7	7,5	9	6	8,7	8,8
3. mondat	7	7	6,6	7,7	7,5	7,6	8,7
4. mondat	6,7	6,9	6,6	6,8	7,6	8,2	10,3
5. mondat	7	6,2	6,3	7	12	7,2	8,8
6. mondat	8,8	6	8,3	7,5	7	9,6	9,1



3. ábra

Az artikulációs tempó alakulása a 6 példamondatban

A táblázat számadatai (időegység/hang) valamint a 3. ábra térbeli görbéi közvetve a mondatok egyes szavainak a kiejtési tempóját mutatják. A kisebb szám esetén a tempó gyorsabb, a nagyobb szám esetén pedig lassabb. Az adatokból



látszik, hogy a vizsgált mondatok közepén (3., 4., 5. szó) a tempó általában gyorsabb, mint a mondat elején, illetve végén. Ezt elfogadhatjuk egy általános szabálynak. Egyes kiugró adatoknál megnéztük a várt tempótól való eltérés okát. Például a 2. mondat 4. szavában mért 9-es érték (igen lelassul a tempó) azzal magyarázható, hogy a szó nyomatékkaal volt kiejtve. Hasonló a helyzet az 5. mondat 5. szavával is.

## Hangidőtartamok

A 20 mondatnyi anyag feldolgozásával nem mindig jutottunk annyi adathoz, amennyire a statisztikaszzerű feldolgozáshoz szükség lenne. Ezért csak azokról a hangokról közlünk adatot, amelyek kellő számban előfordultak (minimum 20) a vizsgált szövegekben (a vizsgálat folytatódik, az adatbázist bővítjük).

A 20 mondat magánhangzóinak (304), valamint néhány mássalhangzó feldolgozásából a következő adatokat kaptuk.

A magánhangzó-időtartamok eloszlását a 3. táblázat mutatja. A táblázatban szereplő számok azt mutatják, hogy adott magánhangzó (1. oszlop) adott hanghosszúságban (1. sor) hány esetben fordult elő (4. ábra).

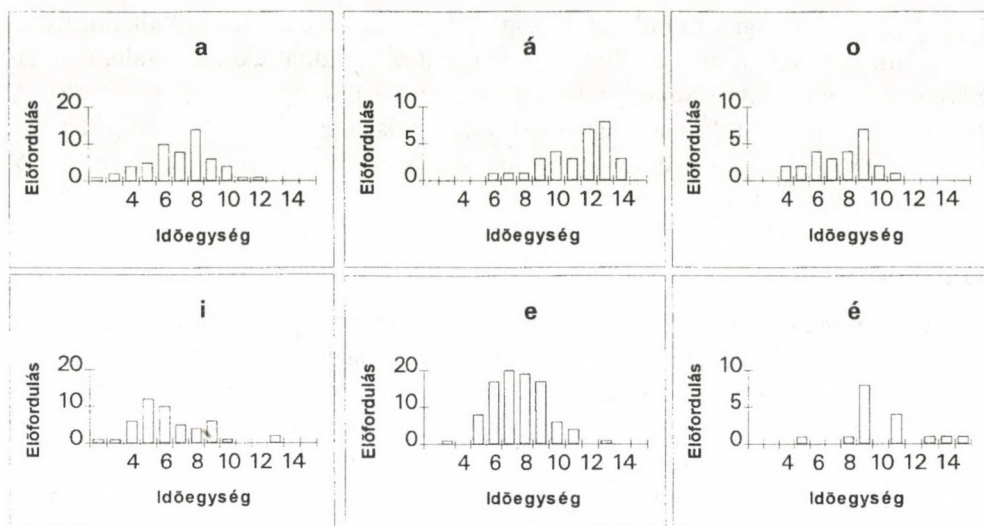
3. táblázat

A hanghosszúság elemi időegységgel kifejezve														
	3	4	5	6	7	8	9	10	11	12	13	14	15	16
hang	Az előfordulások száma													
a	1	2	4	5	10	8	14	6	4	1	1			
á					1	1	1	3	4	3	7	8	3	
o			2	2	4	3	4	7	2	1				
i	1	1	6	12	10	5	4	6	1			2		
e		1		8	17	20	19	17	6	4		1		
é				1			1	8		4		1	1	1

Látható, hogy a **legegységesebb hanghosszúság az e hangnál tapasztalható**. Az eloszlás magja a 7, 8, 9, 10 időegységnyi hanghosszúságoknál található, ami 61 ms-tól 88 ms-ig terjedő hangidőtartamot jelent. A vizsgált *e* hangok 77%-a ebbe a sávba esik.

Erős ingadozás mutatkozik az *á*, *a*, *é* és *i* hangoknál.

A legnagyobb az ingadozás az *a* hangnál, az átfedés közel 7-szeres, azonban ha a kirívóan rövid és hosszú egyszeri előfordulásokat nem vesszük figyelembe, hanem csak az eloszlási adatok maximumát és az ahhoz közeli értékeket, akkor a jellemző időtartam 7-9 időegység, ami 60–80 ms-nak felel meg.



4. ábra

A hanghosszúságok eloszlása a vizsgált magánhangzókra

Az **á** esetében a legkisebb hanghosszúság 7 időegység, a legnagyobb 20 (ezt a táblázatban nem tüntettük már fel). Az átfedés tehát közel háromszoros. A leggyakrabban előfordult hosszúság a 13, 14 időegységre esik, ami 114, 123 ms-nak felel meg. Érdekes figyelmet szentelni arra, hogy az **á** (mint hosszú kategóriájú hang) eloszlásának alsó része (6-12 időegység) átfedésben van az **a** (mint rövid kategóriájú hang) maximum feletti részével, ami azt jelenti, hogy az **á-k** 30%-a időtartamban közel megegyezik az **a-k** közel 40%-ával. Ez azt mutatja, hogy a hanghosszúságok beállításánál esetenként a hosszú hangot rövidítenünk kell, a rövidet pedig nyújtanunk. Az erre vonatkozó általános szabály megtalálása még további vizsgálatokat igényel.

Kisebb az ingadozás az **é** hangnál, ahol az átfedés mintegy 2,5-szeres, a legrövidebb és a leghosszabb realizáció között. A jellemző érték 10 egység, ami 88 ms-nak felel meg. **A két vizsgált hosszú magánhangzóból tehát az á lényegesen hosszabb, mint az é.** Érdekes felfigyelni arra is, hogy az **é** hang annyira rövid a vizsgált mondatokban, hogy jellemző hosszúsága nagyon közel kerül a rövid kategóriába tartozó **e** hang hosszúságához.

Az egyes magánhangzók átlagolt időtartama a 20 mondat alapján a következő:

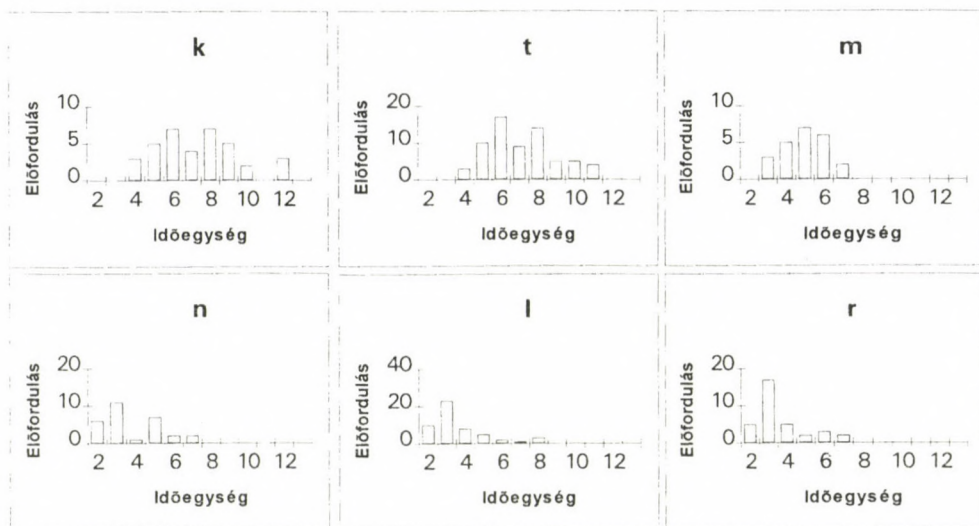
$$\acute{a} = 120 \text{ ms}, a = 80 \text{ ms}, o = 90 \text{ ms}, i = 55 \text{ ms}, \acute{e} = 90 \text{ ms}, e = 75 \text{ ms}$$

A mássalhangzók közül a **t, k, m, n, l, r** hangokra kaptunk megfelelő számú adatot. A feldolgozás eredménye a 4. táblázatban látható.



4. táblázat

A hanghosszúság elemi időegységgel kifejezve															
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
hang	Az előfordulások száma														
k			3	5	7	4	7	5	2		3				
t			3	10	17	9	14	5	5	4					
m		3	5	7	6	2									
n	6	11	1	7	2	2									
l	10	23	8	5	2	1	3								
r	5	17	5	2	3	2									



5. ábra

A hanghosszúságok eloszlása a vizsgált mássalhangzókra

Az adatokból látható, hogy a **k**, **t** zárhangok jellemző hosszúsága 6–8 időegységnyi, azaz 53–70 ms, de szép számban előfordulnak mind rövidebb, mind pedig hosszabb realizációk. Az eloszlási görbe a **k**, **t** hangokra tehát eléggé lapos. Ez a jellemző adat hangsorbelseji helyzetre vonatkozik, ugyanezen hangok abszolút hangsorvégi helyzetben közel kétszeres időtartammal realizálódnak.

A **m**, **n** hangok esetében éles időtartamkülönbség látható a két hang között. A **m**-nél az eloszlás lapos görbét mutat, amelynek maximuma a 4–6 időegységre esik. Ez jellemzően 35–52 ms-os hanghosszúság. Az **n** hang esetében a maximum sokkal élesebb és a jellemző hanghosszúság a 3-as időegységnél van, ami 26 ms-ot jelent. A két hang között tehát lényeges időtartamkülönbség van, amit célszerű kihasználni a mesterséges beszédelőállításnál, ugyanis a gépi beszédet

előállító rendszerekben mindig problémát okoz a *m* és *n* hangok megfelelő hangzásának megtalálása és beállítása.

A *l*, *r* hangok méréseiből egyértelműen látszik, hogy a legtöbb előfordulás a 3-as időegységhez tartozik, ami 26 ms-os jellemző hangidőtartamot jelent mindkét hangnál, holott korábbi mérésekből az olvasható ki, hogy a *l* hang közel kétszer olyan hosszú, mint a *r*.

### A mérési eredmények összevetése korábbi adatokkal

Jelen vizsgálat eredményeit összevetve Magdics (1968), Kassai (1979), Szende (1976) mért adataival látható, hogy a beszédtempó tekintetében a 14,5 hang/s-os érték lényeges tempónövekedést jelent például a Magdics által mért 11,35-ös értékhez képest, ami megegyezik Kassai legutóbbi vizsgálatának eredményével (ebben a kötetben). A legszembetűnőbb hangrövidülések mértékei a következők:

$$\acute{a} = 15\%, \acute{e} = 25\%, k, t = 40\%, m = 45\%, n = 55\%, l = 35\%$$

A rövidüléseknek még nagyobb súlyt ad az a tény, hogy Szende és Kassai adatai spontán beszéd méréseiből kerültek ki, a jelenlegi eredmények pedig olvasott anyagból származnak.

### Összefoglalás

Az eddigi mérésekből három szabály fogalmazható meg a beszéd időszerkezetére vonatkozóan.

1. A beszédhangok szintjén meg tudjuk adni – adott beszédtempóra vetítve –, hogy a magánhangzók milyen hosszúságértékeket milyen valószínűséggel vehetnek fel.

2. Meghatározott beszédsebességhez meg tudjuk adni, hogy a különböző szótagszámú szavak alapbeállításban milyen hosszúsági kategóriába kell, hogy essenek. Ezzel valamelyest irányítani lehet a mondat szavainak tempóbeállítását.

3. A mondat elején, illetve végén meg tudjuk adni, hogy az alapbeállításához képest milyen mértékben kell a szavak tempóját lassítani, hogy közelítsünk a természetes ejtés ritmikai szerkezetéhez.

### Irodalom

- Collier, R. – Leeuwen, H. C. – Willems, L.: Speech synthesis today and tomorrow. Philips J. Res. 47. 1992. 15–3
- Kassai Ilona: Időtartam és kvantitás a magyar nyelvben. NytudÉrt 102. Budapest 1979.
- Magdics Klára: A magyar beszédhangok időtartama. NytudÉrt 67. Budapest 1969. 4.
- OM 8210 Users Manual. Philips. 1988.
- Olaszy Gábor: A magyar beszéd leggyakoribb hangsorépítő elemeinek szerkezete és szintézise. NytudÉrt 121. Budapest 1985.



Olaszy Gábor: Elektronikus beszédelőállítás. Budapest 1989.

Olaszy, G. – Németh, G. – Gordos, G.: The MULTIVOX multilingual text-to-speech converter. In: Talking Machines: Theories, Models and Designs ed. G. Bailly, C. Benoit and T. R. Sawallis. Amsterdam 1992.

Szende Tamás: A beszéd folyamat alaptényezői. Budapest 1976.

**A tanulmány alapját képező kutatás az OTKA 3211 támogatásával történt.**

# BESZÉDKÉPZÉST SEGÍTŐ VIZUÁLIS GÉPI MÓDSZER

**Vicsi Klára, Vassné Kovács Emőke\*, Barczikay Péter**

**MTA Békésy György Akusztikai Kutatólaboratórium**

**\*Bárczy Gusztáv Gyógypedagógiai Tanárképző Főiskola**

## **1. A kutatás célja**

Vizuális oktató és gyakorló rendszer létrehozását tűztük ki célul, amely nagyothalló, siket, ill. beszédhibás gyerekek és felnőttek helyes artikulációjának kialakításához nyújt segítséget. Ez a rendszer főként szurdopedagógusok, logopédusok és foniáter orvosok munkáját teszi hatékonyabbá a helyes ejtés kialakításában, és begyakorlásában. A koncepció része az is, hogy már a gyermekek is önállóan dolgozhassanak, játékos programokkal kötve le figyelmüket.

## **2. Az új módszer alapja**

A beszédképzéskor keletkező hang jellemző paraméterei, a színek, az alaphangváltozás, a hangosság stb. a számítógép képernyőjén ma már azonnal, grafikus és igen szemléletes módon jeleníthetők meg. A képernyőn megjelenített képet nevezzük 'hangképnek'. Amennyiben jó a jellemző paraméterek meghatározása és a grafikai megjelenítés, a hibás képzéskor megjelenő 'hangképek' erősen eltérnek a helyes ejtést jellemző 'hangképek'-től. Így lehetőség nyílik arra, hogy a kísérleti személy, látva a saját rossz hangmintasorát és annak eltérését a normál ejtés hangmintasorától, hatékonyan rávezethető a helyes hangsor előállítására. Nagyothallók és siketek esetében ez a vizuális visszacsatolás helyettesíti az akusztikai visszacsatolás hiányosságait (1. ábra).

A helyes artikuláció kialakítása minden esetben új szenzomotoros integráció kiépítését kívánja meg. Ép és hallássérült egyéneknél egyaránt a multiszenzoros megközelítést követi a terápia, a hallás, a látás, és újabban a tapintás bekapcsolásával.

A hallássérülteken kívül a centrálisan károsodottaknál (pl. dysgnoziás, dyspraxiás figyelem és emlékezet zavaraiiban szenvedők), hasadékosok, többszörösen beszéd-sérült egyéneknél azonban a korrekt szenzoros és motoros fejlesztés ellenére az integráció kialakulása igen nehézkes. Vagyis a rögzítés és az automatizálás fázisa elhúzódik. Tehát minden olyan eljárás, eszköz, amely ezt az időszakot lerövidíti, megkönnyíti a szakember munkáját.



### 3. A kutatás készültségi szintje

Célunk az, hogy a módszert a teljes hangrendszerre kidolgozzuk, az életkort, valamint a különböző etiológiai hátteret figyelembe véve. A Soros Alapítvány támogatásával 1991. júl. 1. és 1992. jún. 31. között végeztük a kutatást. Ez alatt az egy év alatt a sziszegők korrekciójára alkottunk meg egy rendszert, amelyet a későbbiekben ismertetünk. Szurdopedagógusok, logopédusok, foniáter orvosok számára azt kívántuk bebizonyítani, hogy ez az újfajta módszer alkalmas kitűzött célunk megvalósítására és érdemes a kutatásával, a rendszer kibővítésével a továbbiakban is foglalkozni.

### 4. A rendszer felépítése

A rendszer bemenete a hallást leegyszerűsítve modellező műszer, amely IBM személyi számítógépbe behelyezhető egység. Ennek a speciális preprocesszornak a segítségével (Zwicker 1982) közelítőleg a beszédhangok olyan jellemzőit tudjuk mérni, amelyek a belső fülben keletkeznek (cohcogramok). A mikrofonba mondott szavak vagy mondatok feldolgozott akusztikai képe a számítógép képernyőjén jelenik meg (2. ábra). A rendszer a beszédhang lényeges paramétereit emeli ki, és illusztratív, játékos módon teszi láthatóvá. A lényegkiemelésen kívül, szóeleje-vége meghatározást, dinamikus és lineáris idővetemítést, távolságszámítást stb. alkalmaztunk mint speciális műveleteket (Vicsi 1986).

### 5. A kutatás rövid összefoglalása

Munkánkat a sziszegőkkel, azaz a réshangok és az affrikáták egy csoportjával [s, z, ʃ, ʒ, tʃ, tʃ̥] kezdtük, mivel ez a hangcsoport a legsérülékenyebb, a terápiában pedig a legmakacsabb. Ennek okai a következők: a legbonyolultabb képzési hangok közé tartoznak, relatíve szűk artikulációs helyen képződnek, a képzést erősen befolyásolják a szájtéri viszonyok, a hallás épsége, a hallási percepció szintje, a beszédszervek motoros képessége.

A sziszegők kóros ejtését szigmatizmusnak nevezzük, a pöszeségen belüli megoszlása az előbb említettek miatt egyre növekvő tendenciájú.

#### 5.1. A korrekt ejtés hangképi adatainak meghatározása

A sziszegő hangok színképi paramétereinek statisztikai vizsgálatát végeztük el: az átlagértékek és a megengedett szórás meghatározásával. A hibás ejtést ezekhez az átlagértékekhez viszonyítottuk.

## 5.2. Hangkép torzulásának vizsgálata

Különböző típusú beszédzavarok esetén a beszéd hangképe az adott eltérés típusára jellemző. Ezt a jelenséget hallássérült és éphalló egyéneknél a következő szigmatizmusok esetében vizsgáltuk: interdentális, stridens, laterális. Itt támaszkodni tudtunk a korábbi kutatási eredményekre is (Vicsi – Lugosi – Linder 1987).

## 5.3. A korrekt ejtési mintatár létrehozása

A logopédiai, szurdo-logopédiai gyakorlatban alkalmazott hangsorok, tipikus ejtészvariációk szavak mondatok szűkített mintatárát hoztuk létre. A rögzítés és az automatizálás fázisában u.i. a beszédhibás egyén a saját hangképét a korrekt ejtési mintatár hangképeivel veti össze.

## 5.4. Idővetemítési optimalizálás

A korrekt ejtés és a hibás ejtés hangképének összehasonlításakor az időbeli eltérések kompenzálására különböző idővetemítő eljárásokat próbáltunk ki. Erre azért van szükség, mert mindenki más tempóban beszél. A hangkép viszont vizuálisan akkor hasonlítható jól össze, ha megjelenítéskor az időbeli eltérést kiküszöböljük. Más-más vetemítést kell használni nagyothallók és éphallásúak esetében, valamint a terápia egyes fázisaiban.

## 6. A terápia módszertana

Először, a képernyőre való figyelem irányítását, a beszédszervi mozgások megjelenítésének, változásának kontrollálását kívánjuk elérni. A különböző hangokhoz ugyanis más és más jellegzetes hangkép tartozik. Annak ellenére, hogy egy új módszert vezetünk be, a hagyományos logopédiai terápia jól bevált lépéseit követjük: előkészítés, hangfejlesztés, rögzítés és automatizálás.

A hang ELŐKÉSZÍTÉSÉ-hez különböző nyelvgyakorlatokat használunk, pl. *tyu-tyu, lá-lá* stb. (3., 4. ábra). A folyamatos levegőáramlás gyakorlására a képen látható holdat vagy napot kell elérni a színek egyes összetevőinek (5., 6. ábra).

HANGFEJLESZTÉS-nél először a hangképzőszervek helyes elhelyezkedését, a hang képi szimbólumával, valamint a megfelelő grafémával együtt mutatjuk be a gyerekeknek. A rendszerben kidolgozott hat hang betűjele a következő: *sz, z, s, zs, c, cs* (7., 8., 9., 10., 11., 12. ábra).

A helyes kitartott hang színekpi összetevőinek átlagát és szórását játékos formában jelenítettük meg a képernyőn, pl. a [s] hang esetében a sárga út elválasztó vonala jelzi a normál ejtés átlagértékét, az út maga a megengedett szórást. Korrekt kiejtésben a megjelenő hangkép a sárga úton belül kell hogy maradjon, a sárga út vonalát követve (13. ábra). A [z] hangnál a méhecske útján belül kell



tartózkodni (14. ábra). A [j] hangnál a folyó nyomvonalát kell követni (15. ábra), a [z] hangnál az időalagútban kell a repülőnek haladni (16. ábra).

A RÖGZÍTÉS-ben a gyermek a normál ejtés színeképek időbeli változásait, mint összetett hangképet látja. Szaknyelven ez a 'COCHLEOGRAM' (17. ábra). A képernyő felső felében jelenik meg a szótag színeképi etalonja, alatta pedig a gyermek által produkált ejtés.

A helyes képzés kontrollálását rajzosan segítjük elő. Például a [s] hang esetében rossz az ejtés, ha a kígyó fejét eltakarják a színeképi összetevők, vagyis a korongok. A [j] hang esetében korrekt ejtésnél a vonat füstjét el kell takarni, magának a vonatnak látszania kell. Hasonlóan járunk el a többi hangnál is. Ilyen módon szoktatjuk rá a gyereket arra, hogy a hangképből a lényeges részekre figyeljen. A sziszegő hangokat sárga/piros színárnyalatokkal emeljük ki a képből. A gyakorlás során így a gyerek könnyen elsajátítja a színekép olvasásának képességét. Könnyen észreveszi a hibás ejtést, és azonnal korrigálni tudja azt. A rögzítés fázisában az újonnan megtanult hangot szótagokban gyakoroljuk, először a szótag elején, végén, majd szótagközi helyzetben (18. ábra).

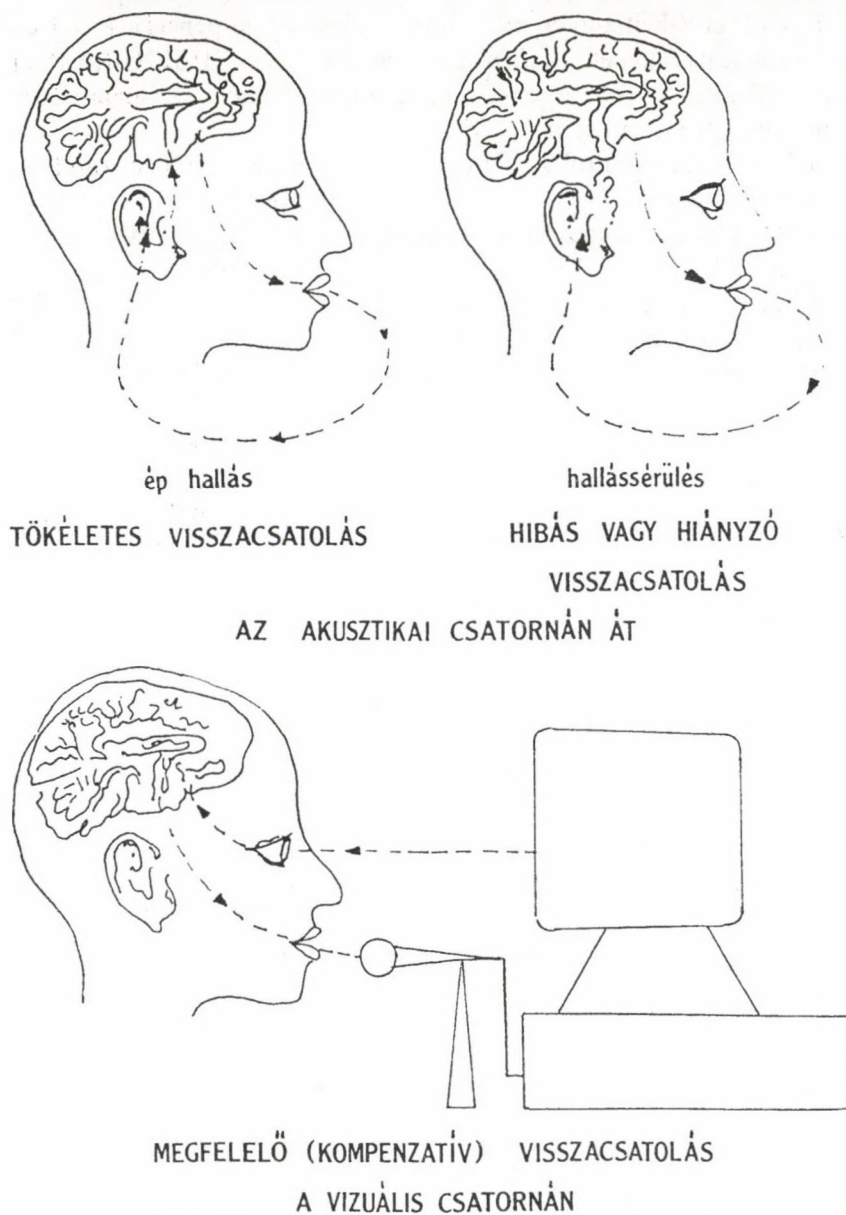
Az AUTOMATIZÁLÁS-nál szavakba, majd mondatokba építjük be az új hangot a már említett módon.

## 7. Eredmények

A módszer kipróbálására és a módszertan kidolgozására nagyon rövid idő állt rendelkezésre. Ezért a módszer fejlesztése és a terápia egyidejűleg folyt. Ez a megoldás igen hatékonynak bizonyult, mert a szükségletnek megfelelően tudtuk a módszert alakítani. A módszer alkalmazhatóságát éphalló és hallássérült gyermekeken próbáltuk ki.

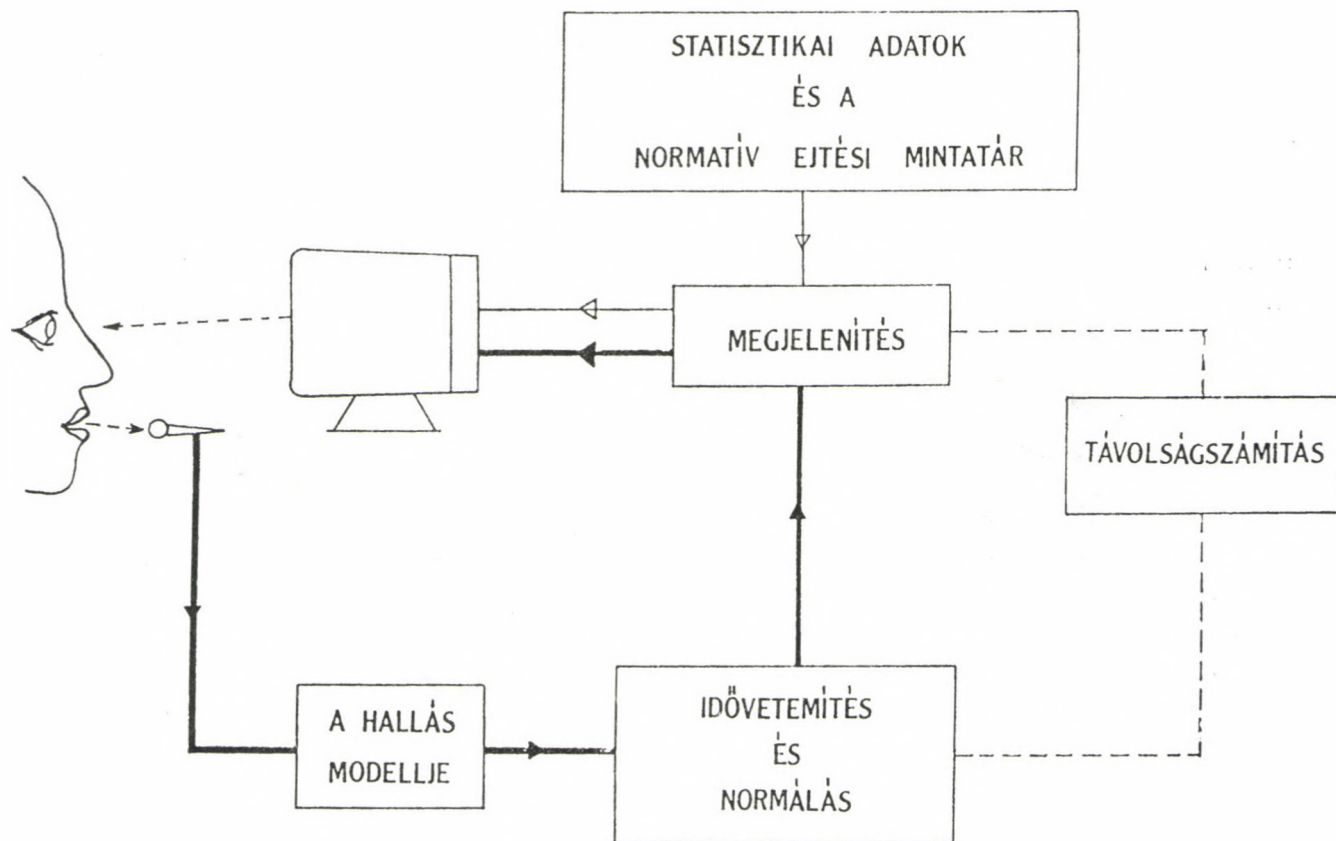
Összegezve elmondhatjuk, hogy a gyermekek ezzel a módszerrel igen jól motiválhatóak, figyelmük jobban összpontosítható, az egyes hangok kialakítása rövidebb időt vesz igénybe, mint a hagyományos terápiával. Súlyos hallássérülteknél hagyományos szurdopedagógiai módszerrel sokszor nem is lehet kialakítani, ill. rögzíteni a kívánt hangot. Ez leginkább a sziszegőkre érvényes. Ezzel az új vizuális visszacsatoláson alapuló módszerrel viszont rövid idő alatt sikerült egy gyakorlatilag süket 11 éves gyermek [j] és [s] hangját helyesen kifejleszteni. Eredményeinkről nemzetközi fórumokon is beszámoltunk (Vicsi et al. 1992a, 1992b).

Eljárásunkat a magyar hangrendszerből csak hat hangra tudtuk ilyen rövid idő alatt kidolgozni. Kutatásunkat a jövőben a többi magyar beszédhangra is ki kívánjuk terjeszteni.

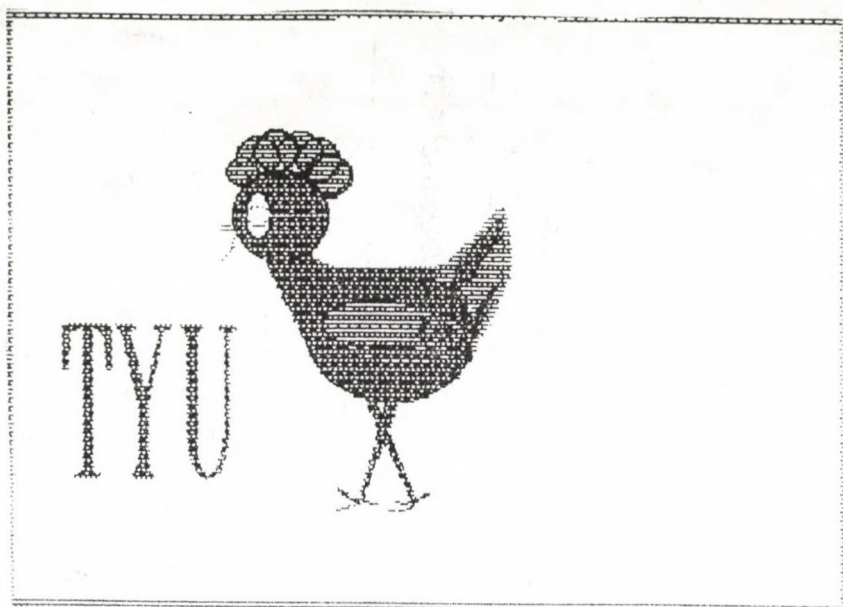


1. ábra  
Általános koncepció

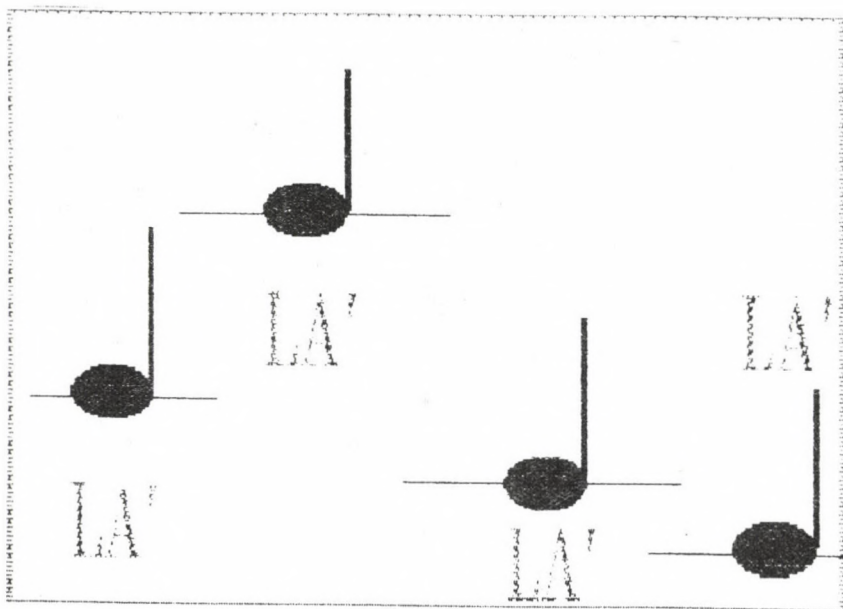




2. ábra  
A vizuális oktatórendszer felépítése

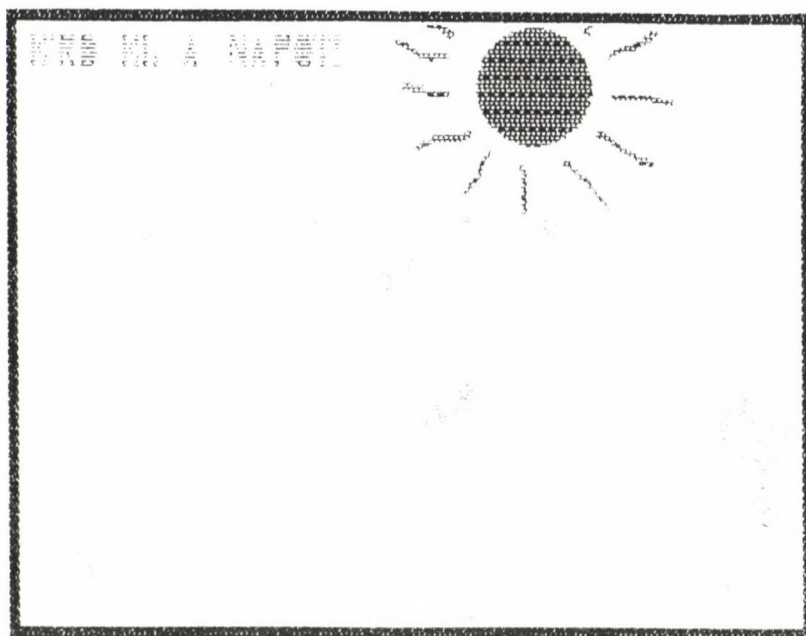


3. ábra

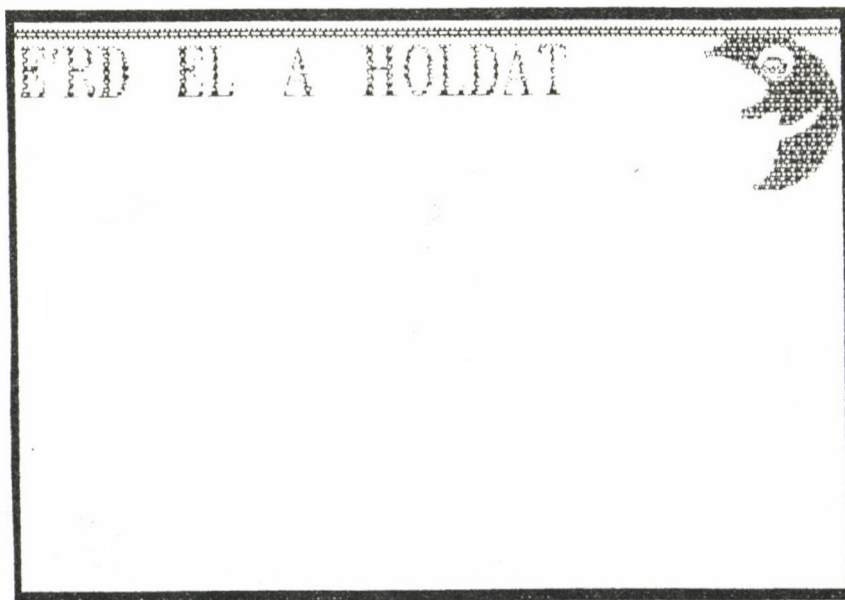


4. ábra

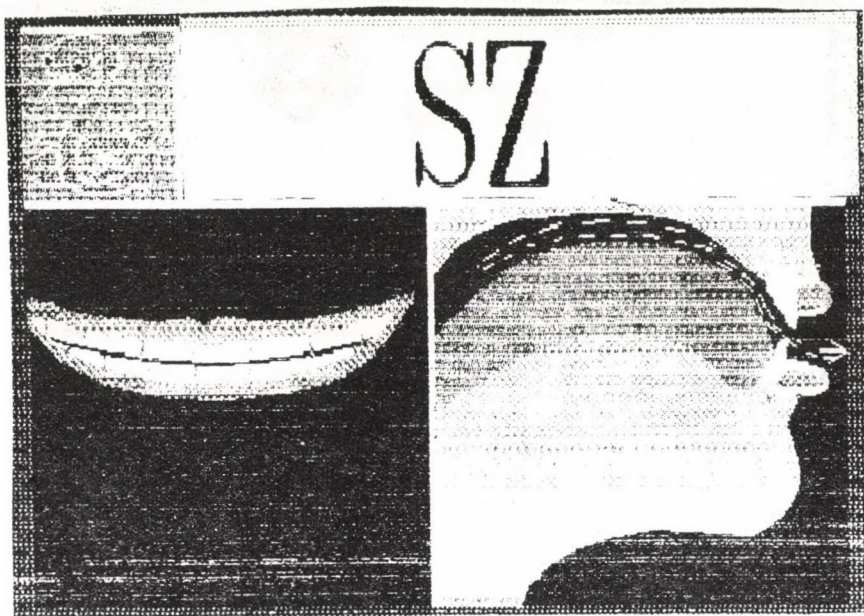




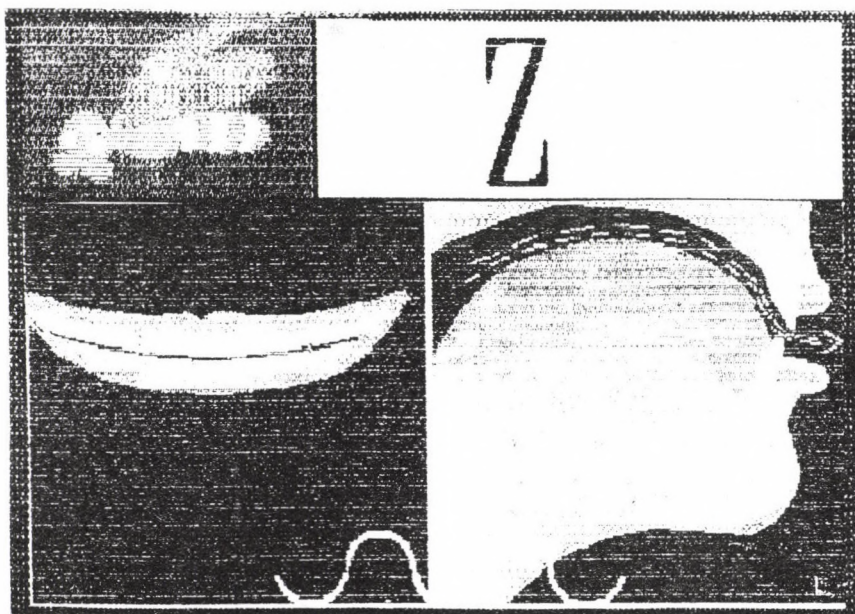
5. ábra



6. ábra

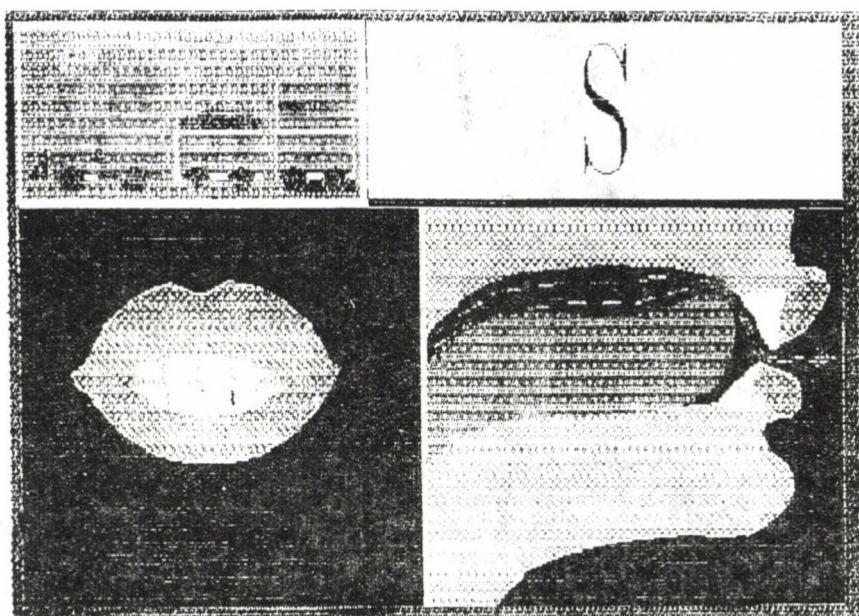


7. ábra

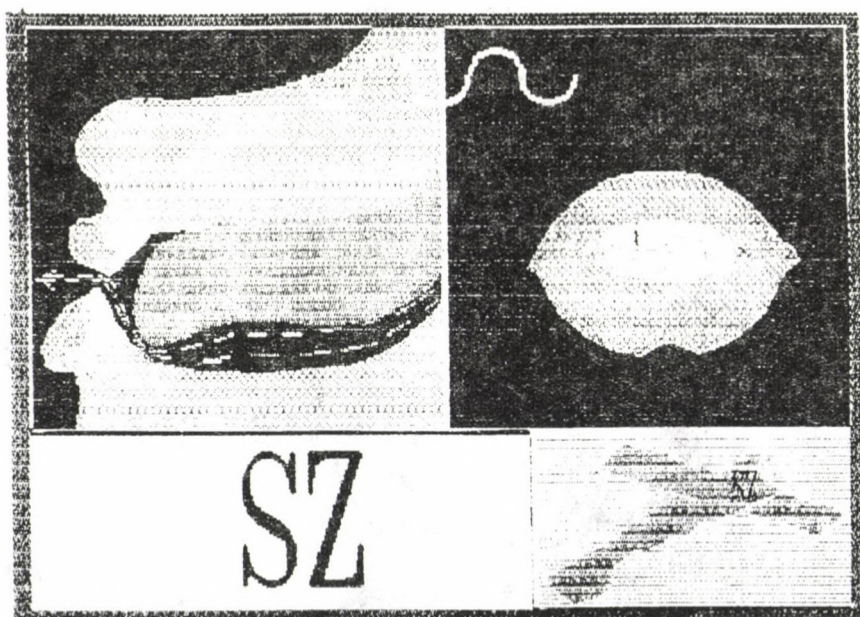


8. ábra



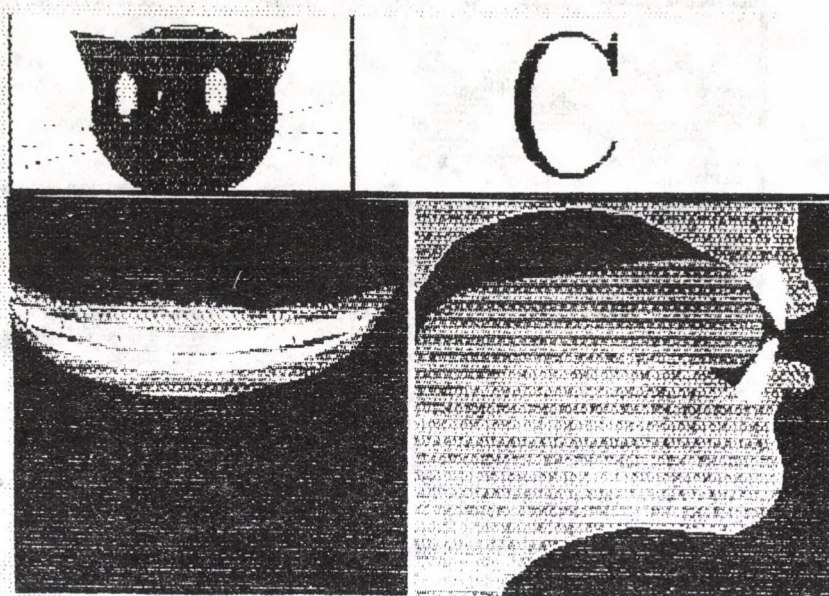


9. ábra

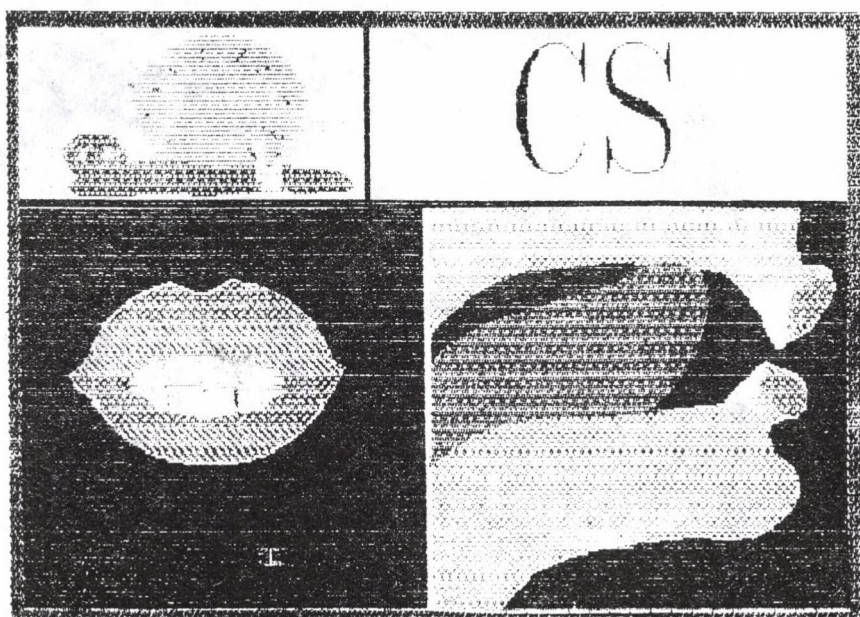


10. ábra





11. ábra

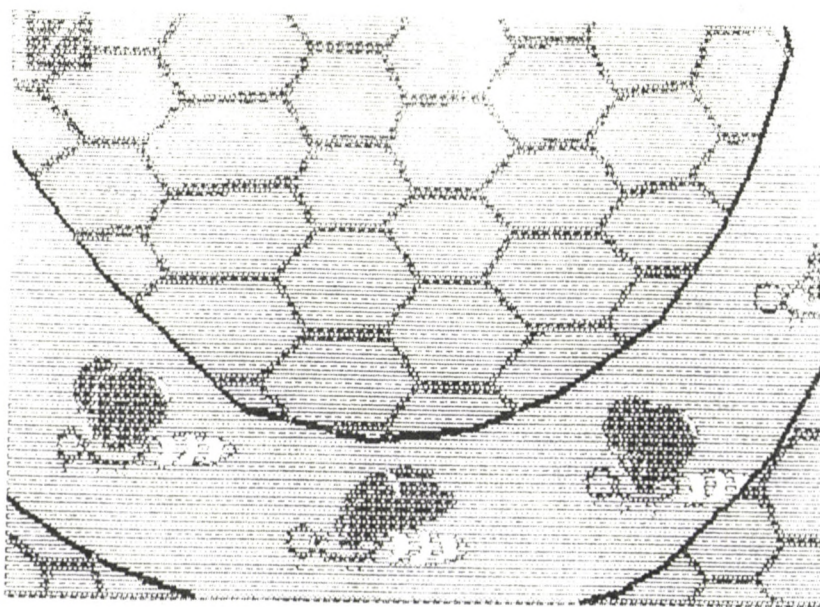


12. ábra



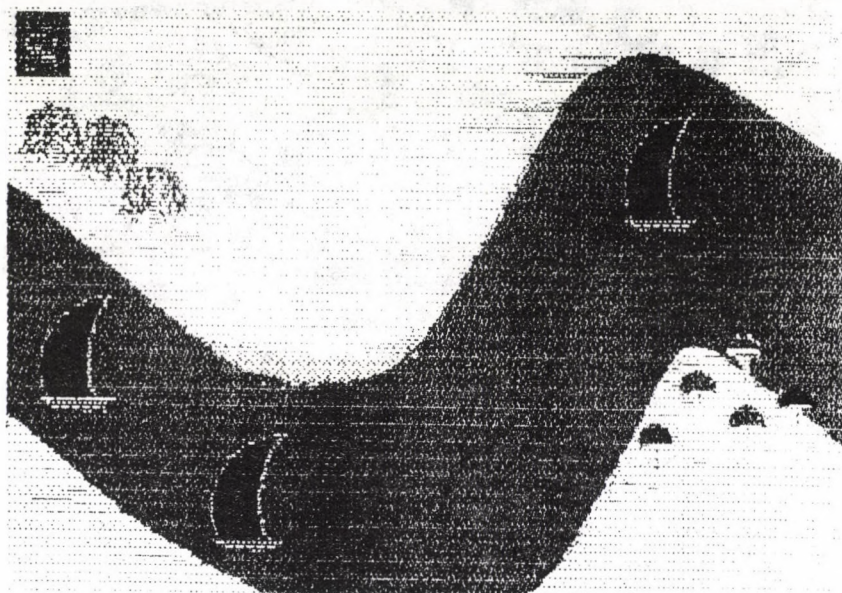


13. ábra

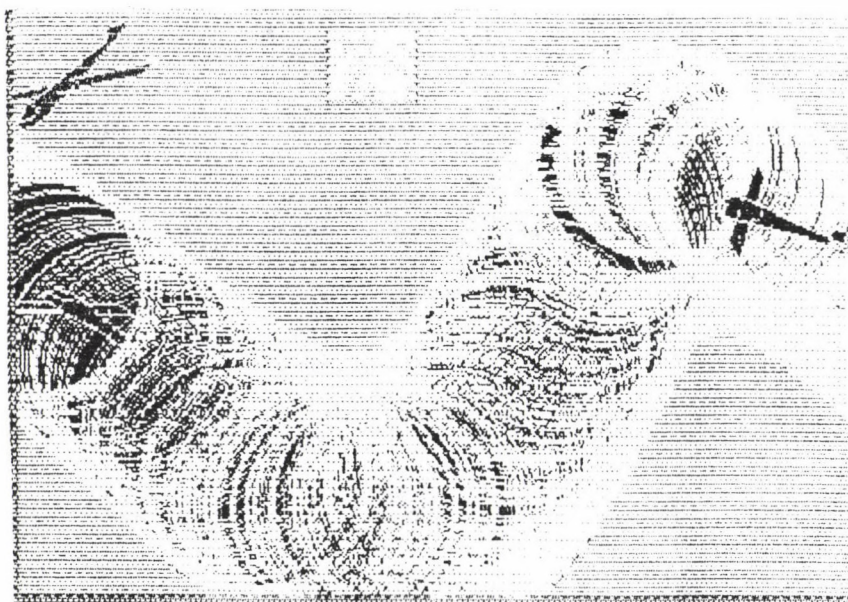


14. ábra





15. ábra

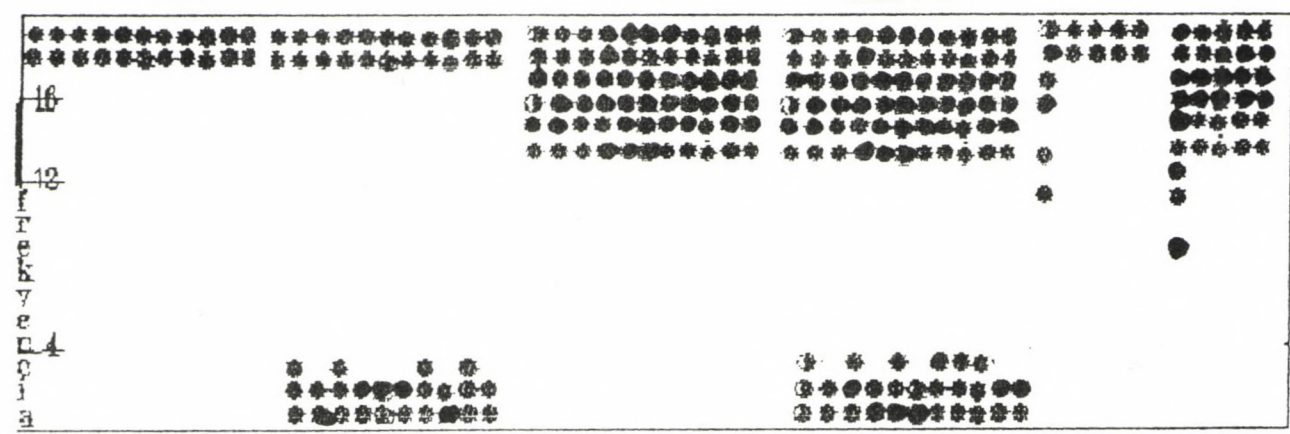


16. ábra



# MINTACOCHLEOGRAMOK

SZ      Z      S      ZS      C      CS



Időtartam

17. ábra



18. ábra



**Irodalom**

- Vicsi, K.: Automatic Word Recognition System Using Some Results of Psychoacoustics. 6. FASE Conference, Sopron 1986. 3. A. 8.
- Vicsi, K., Lugosi, T. and Linder, T.: Search for Optimal Teaching Procedure and Warping Algorithms for Isolated Speech Recognizers. Proc. XIth IcPhS. Tallin 1987. 363.
- Vicsi, K., Kovács-Vass, E., Barczikay, P.: Visual Tools for Speech Handicapped Persons. 3rd. Int. Conf. on Computers for Handicapped Persons. Vienna 1992a, 540.
- Vicsi, K., Kovács-Vass, E., Barczikay, P.: Speech Picture in Speech Therapy. XXII. Word Congress of Logopedics and Phoniatrics. Hannover 1992b.
- Zwicker, E.: Psychoakustik, Springer Verlag, Berlin 1982.

# NYELVI STRUKTÚRÁK HELYREÁLLÍTÁSA GLOBÁLIS AFÁZIÁS BETEGEK KEZELÉSÉBEN

Hegyi Ágnes

Szent-Györgyi Albert Orvostudományi Egyetem  
Ideg- és Elmegyógyászati Klinikája, Szeged

## 1. A terápia elméleti alapjai

**1.1.** Az afázia a nyelvi képesség zavara. Nyelvi képesség: Az ember mint magasabb rendű élőlény anatómiai és neurofiziológiai adottságából eredően rendelkezik olyan más élőlénytől megkülönböztető képességgel, amely – a II. jelzőrendszeren keresztül – lehetővé teszi a külvilágból érkező információk közvetítését és megértését oly módon, hogy a közösség által elfogadott verbális jeleket a nyelvtan elfogadott szabályai szerint használja, kiegészítve őket a nem verbális jelekkel. A nyelvi képesség nemcsak feltétele a sikeres nyelvi tevékenységnek, hanem maga is a tevékenység során fejlődik.

A nyelvi képességről alkotott definíciónk egyértelműen meghatározza az afáziaterápia célját. A nyelvi struktúrák helyreállítása során célunk, hogy az afáziás beteget újból képessé tegyük a külvilágból érkező információk megértésére és közvetítésére. Tehát újból használja a beteg a nyelvi közösség egyezményes verbális jeleit a nyelvtan elfogadott szabályai szerint, kiegészítve őket a nem verbális jelekkel, a nyelv valamennyi modulációjában.

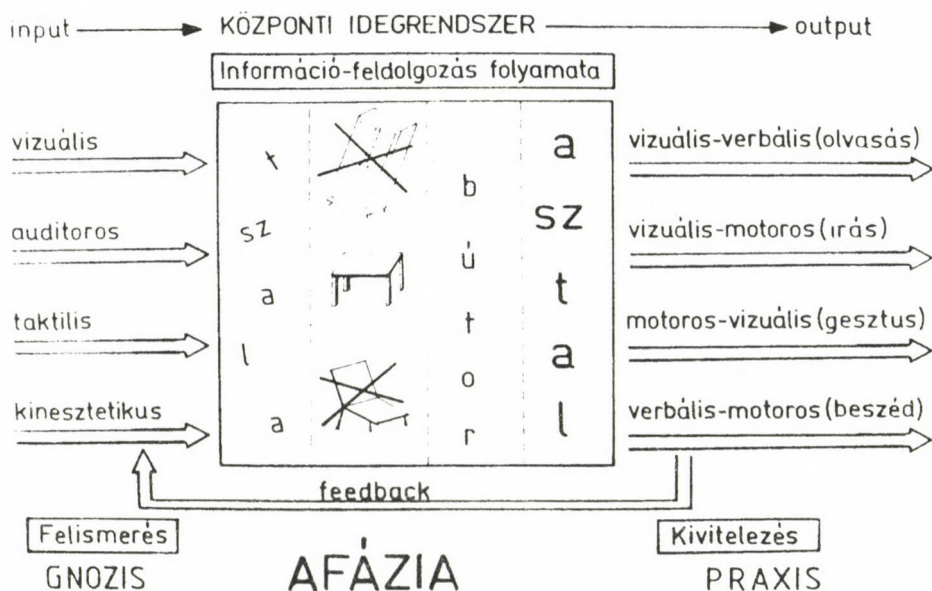
A fentiekből következik, hogy a nyelvi képességet helyreállító terápiáknak módszerükben különbözniük kell a hangképzési zavaroknál és a beszédzavaroknál alkalmazott terápiás eljárásoktól (Lundman – Tenenholtz – Galyas 1978, Kassai 1982, Hegyi 1988).

**1.2.** Ha az afáziák közös problémájára szeretnénk magyarázatot kapni – definíciónkból kiindulva – egyrészt el kell határolnunk a nyelvi képesség más zavaraitól. Másrészt értelmeznünk kell az afáziát és az afáziával szorosan összefüggő, ám attól megkülönböztethető más funkciók zavarait: a felismerő funkciót, a gnóvizst; a kivitelezést, a praxist és annak különböző formáit az írást, az olvasást, a beszédet és a gesztust (1. ábra).

A kommunikációs modell bemenő oldalán (input) az egymást erősítő ingerek az elsődleges érzékelési csatornán keresztül érintik a központi idegrendszert. A különböző ingerek percepciója után megtörténhet az információ felismerése és értelmezése. Az információ nyelvi feldolgozásának modelljei szerint az értelmezés folyamata feltételezi az adekvát szemantikai, fonológiai, fonetikai illetőleg grafikai reprezentációt. Ezek együttesen a kimenő oldalon (output) a kivitelezést



több csatornán keresztül teszik lehetővé. A visszacsatolás ellenőrző folyamata meggyőzhet a kivitelezés helyességéről, amely olvasás, írás, beszéd vagy a sike-  
teknél a beszédet helyettesítő "kézjelnyelv" (Poizner – Beluggi – Klima 1990) és  
gesztuskommunikáció útján is megvalósulhat.



1. ábra  
Információ-feldolgozás a szófelidézés szintjén

**1.3.** Az afáziás beteg hiányos vagy hiányzó nyelvi képességének az okát a központi idegrendszer információ-feldolgozó folyamatának hiányában, illetőleg zavarában látjuk (Marshall – Newcombe 1973; Newcombe – Marshall 1981.). Az információ-feldolgozó folyamat valamely pontján beálló zavar a beteget megakadályozza abban, hogy a nyelvi jeleket a nyelv szabályainak megfelelően előhívja és szándéka szerint a kifejezés eszközeként felhasználja (Finkelnburg 1870; Froment 1921; Head 1926; Helm-Estabrooks et al. 1981; Howard et al. 1984). Ezért feladatunk, hogy segítsük – a nyelv valamennyi szintjén – a nyelvi jelek előhívásának módját, hogy közlési szándéka szerint, azok megfelelő sorát alkossa meg a beteg.

**1.4.** A jelhasználat zavara olvasásban és írásban ugyanolyan áttételesen jelentkezik, mint a beszéd és a beszédmegértés zavarában (Poeck 1982; Lang – Stockert 1986). Az olvasás és írás útján történő információ-feldolgozó folyamatok ismerete az afáziás beteg nyelvi képességének helyreállításában közvetett segítséget nyújt. Ezért feladatunk az egymást erősítő nyelvi modulációknak – az írásnak, az olvasásnak és a beszédnek – egyidejű helyreállítása, amelyben természetes szerepet kap a beszédet kísérő kifejező taglejtés és kézmozdulat (vö. Shewan – Bandur 1986).

## 2. A betegek kiválasztásának szempontjai

A betegek kiválasztásában a klinikai (neurológiai és belgyógyászati) diagnózisra és a funkcionális próbák eredményeire támaszkodtunk. A neurológiai faktorok mellett a nyelvi, a pszichés és a pszichoszociális faktorokat vettük figyelembe (2. ábra).

Neurológiai szempontok	a betegség természete stabilitás a szervezeti állapotban lokális sérülés főként vaszkuláris történés
Nyelvi szempontok	készség a verbális kontaktusra figyelem a beszédhangra alapvető hangadási képesség ritmus visszaadásának képessége
Pszichológiai szempontok	érintetlen intellektus ép emlékező képesség megfelelő motiváltság mentesség a nagyobb pszichiátriai betegségtől
Pszichoszociális szempontok	családi együttműködés környezeti és szociális támogatás elérhető terápiás központ

2. ábra

A betegek kiválasztásának szempontjai

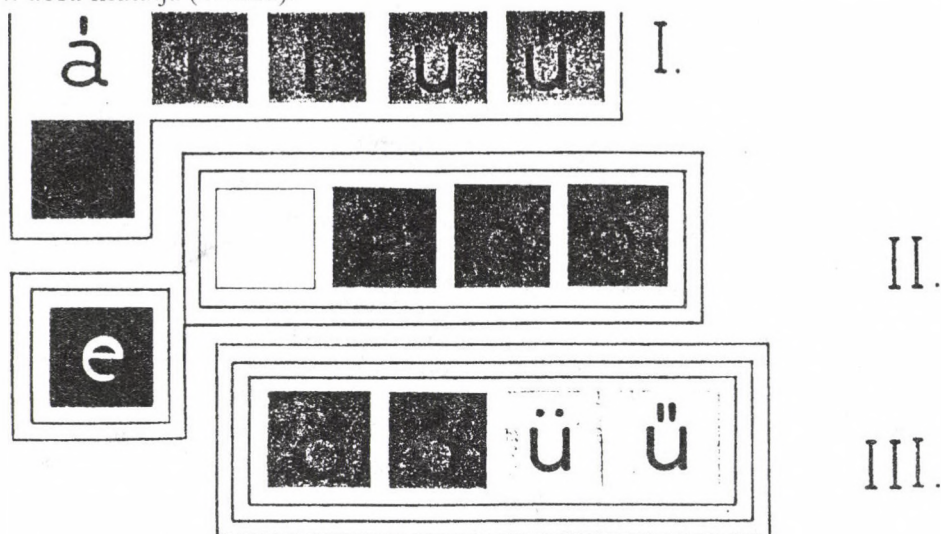




máció-feldolgozó folyamatok működtetése, nyelvi analízis, szintézis és szintézis útján történő analízis segítségével. A terápia "haladó szakaszában" – a mondatalkotás és a szövegalkotás szintjén – a beszédet javítjuk azáltal, hogy az olvasás és az írás útján történő információ-feldolgozás állandó külső ellenőrzést biztosít mind a fonológia, mind a grammatika, mind a szemantika számára, szemben a beszéddel, amelyben az ilyen ellenőrzésnek többnyire belsőnek kell lennie (Lytton – Brust 1989).

Ha az aktivizáló szakaszban a beteg képes előbb színeket, egyszerű mértani alakokat azaz "prototípusokat" (Ádám 1987), majd tárgyakat több tulajdonság alapján is azonos csoportba válogatni, akkor megtartó emlékezete képes több stimulus egyidejű befogadására és megtartására. Ezt a képességet használjuk fel arra, hogy a hangképzéshez kötött nyelvi jeleket a beteg képes legyen felidézni oly módon, hogy egyidejű kapcsolat jöjjön létre a kimondott hang, az elhangzott hang és a leírt betűkép között. – Úgy véljük, a **magánhangzók**, mint zenei elemek felidézését a nyelvhasználat szempontjából nem domináns félteke ép működése is támogatja (Pléh 1981; Lebrun 1983; Geschwind – Galaburda 1985).

A magyar magánhangzók felidezésének sorrendjét három csoportba rendezve a 4. ábra mutatja (4. ábra).



4. ábra

A magánhangzók felidezésének sorrendje

A magánhangzók megkülönböztetését egymástól és azok kiemelését a hang-sorból auditív, vizuális, kinezetikus és taktilis ingerek együttesével segítjük. A hangot kiejtjük, színes alapú betűképét megmutatjuk, majd a képzésmódra utaló vagy a "kódszót" (de Partz 1986) felidéző fonomimikával kísérjük közös hang-



zójtésünket. A hangot kézírással is rögzítjük, majd önálló ejtését szorgalmazzuk.

A **mássalhangzók felidézését** és megkülönböztetését más anyanyelvi hangoktól az első négy vokális biztos felismerése után kezdjük. Itt is, mint a magánhangzókánál az indirekt stimulációs ingeregyüttessel ható felidéző módszert követjük (Weigl 1979).

Már az első négy magánhangzó és egyetlen mássalhangzó felidezésének a képessége biztosítja, hogy a beteg kevés számú hang- és betűkészletéből értelmes szónak megfelelő hang és betűsort alkosson.

**Szóalkotáson** az akusztikusan felfogott szó vizuális megformálását értjük.

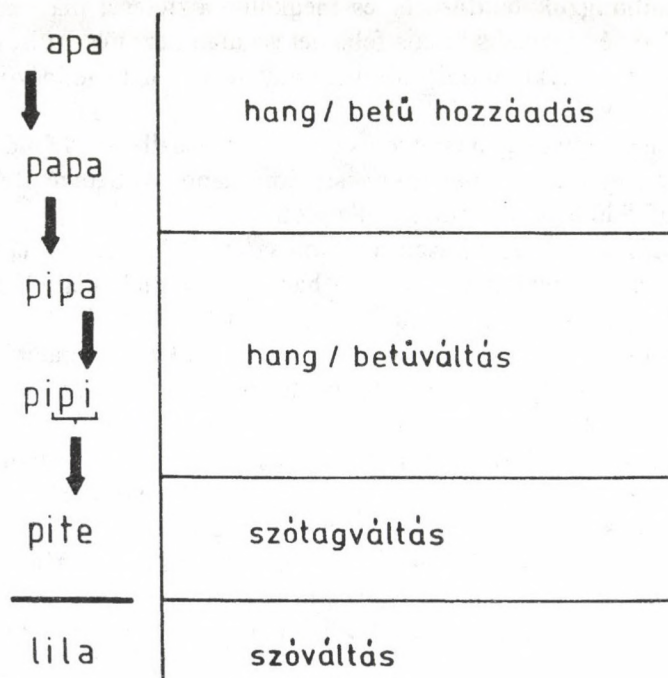
Az akusztikusan elhangzott teljes szó hang- és betűjeleinek előhívása és azok megfelelő hangsorainak megalkotása kezdetben a szavak alaki tulajdonságai szerint alkotott csoportokban történik. A szóalkotást a kezelő stratégiai utasítása segíti, amely utal a hang illetőleg a betűváltozás helyére és arra, hogy milyen változás fog bekövetkezni (5. ábra). Például, *"Figyelve a szó elejére, egy betűvel hosszabb lesz a szó!"* A szó szimbólumértékének felidézését támogatja a hangzás után felismert minimális fonológiai eltérés és a szó szemantikai megjelenítése tárgyi szituációs kontextusban és mondatban (Howard et al. 1985).

A kezelő óráról órára a beteg aktuális hangfelidéző nehézsége szerint választja meg a szavak alaki tulajdonság szerinti csoportját. Ezért nem lehet tartósan eszközünk a szavakhoz tartozó képgyűjtemény a szavak szemantikai tartalmának támogatásában. Ezek helyett a szavakat, például azok etimológiai származékait, mondatokban, a mondatok részeként tálaljuk és ebből a kontextusból emeli ki a kezelő – később önállóan a beteg – a kirakásra váró szót.

A gyakorlatok előrehaladtával, amikor a beteg egyre önállóbb a mondatok által segített szófelidezésben, a szavak jelentéviszonyai alapján alkotott csoportokon belül teljes szóváltással történnek a szóalakítások. A jelentéviszonyok alapján a felidézhető szavak körét a következő fokozatok szerint bővítjük: antonimákkal (*kicsi, nagy*), szinonimákkal (*eb, kutya*), szemantikai mezővel (*bútor, asztal, szék*), rész-egész kapcsolatokkal (*asztal, fiók*), funkcionális kapcsolatokkal (*gyalu, fa, fűrész, enyv*) és szituációs kontextussal (*pénztár, vonat, menetjegy, kalauz, utasellátó, újság, vasútállomás*).

Funkcionális kapcsolatokban és szituációs kontextusban a tárgyak a beteg cselekvéssel kapcsolatos ismeretei alapján jelennek meg. Helyesnek látjuk, ha a gyakorlatokat a **mondatalkotás** szintjén folytatjuk. Ekkor a beszédben kifejezett mondatalkotás az olvasás és az írás közvetítésével valósul meg. Ehhez a diszlexiás gyermekek számára készült Meixner-féle olvasókönyvet használjuk.

Ugyanazon a terápiás órán belül a betűk, a szótagok és a szavak olvasásán keresztül jutunk a mondatok olvasásáig.



5. ábra

Szóalkotás a szavak alaki tulajdonságai szerint

A szavakat szükség szerint, azaz attól függően, hogy a beteg mely olvasási csatornát használja, szelektív korlátozással olvastatjuk. Az olvasott szót megismételtetjük és értelmeztetjük. A szómagyarázatban "verbális" mondatértékű kifejezések és "nem verbális" eszközök használata egyaránt elfogadhatók (Davis, Wilcox 1985).

Az egyszerű, a bővített és az összetett mondatok alkotását mondatanalógiák támogatják, amelyek lehetővé teszik a felszíni jegyek (szórend, ragok, jelek stb.) alapján a tapasztalati (nyelvtani) általánosítást anélkül, hogy a nyelvtani kategóriákat megneveznénk.

Az időrend gyakorlásával a mondatokat eseménysorokká, szöveggé kapcsolja össze a beteg. A **szövegalkotás szintjén** képtörténetek, majd olvasás utáni szövegfelidézés alapján alkot szöveget a beteg.

Friederica Meixner osztrák kolleganő képtörténeteit használjuk (6. ábra).





6. ábra

Képtörténet a szövegalkotás szintjén

A történet egyes képeihez írásban kérdéseket adunk. A kérdésekben a szöveget összetartó elemeket (direktívákat, anafórákat, kötőszavakat stb.) észleli a beteg. A szóban megfogalmazott válaszokban, majd az írásban ezeket kötelezően használja. Használatukkal képes lesz a beteg a megelőző mondathoz igazítani a rá következőt.

Az olvasás utáni szövegfeldolgozás magasabb szinten fejleszti az olvasást és a szövegalkotást. La Fontaine meséinek költői műfordításaiban kevés a redundáns elem, annál több az új információt hordozó kifejezés, amelyek nem támogatják a beteg sztereotip kifejezésesein alapuló kitalálásait (Fónagy 1986). A mese leírásában és elmondásában lehetőség van arra, hogy az idézett párbeszédet a grammatikailag bonyolultabb függő beszéd váltsa fel.

**3.3. A konszolidáló szakaszban** célunk az általános nyelvi kommunikációs képesség gyakorlása a betegcsoport által teremtett szituációs kontextusban. A csoportterápiában irodalmi alkotás; novella, elbeszélés, kisregény elolvasása után egyenrangú kommunikációs partnerként kölcsönös az információcsere. A nyelvhasználatot a társalgási beszéd motiválja (Cvetkova 1980; Glockman 1981).

#### 4. Eredmények és következtetések

1982-től 1990-ig a nyelvi restrukturáló terápiánkban résztvevő afáziás betegekről és eredményeinkről az alábbi táblázat nyújt tájékoztatást (7. ábra).

Betegek száma	19
Nő	7
Férfi	12
Kor szerinti megoszl.	33 – 70 (év)
Átlagéletkor (év) $\pm$ SD	52,9 $\pm$ 12,5
Tanultsági szint	közép és felsőfok
Neurológiai betegség	akut cerebrális lézió vaszkuláris történet (Stroke) bal hemiszférium sérülése
Afázia típusa Kísérő neurológiai tünetek	globális jobb oldali végtagbénulás jobb oldali végtagzsibbadás jobb oldali végtag túlérzékenysége
Átlagos kezelési idő	legalább 4 év azaz 420 óra/fő
Javult	15 fő
Nem javult (4 fő)	alacsony tanultsági szint (analfabéta) alacsony intellektuális képesség ritmus visszakopogásának képtelensége – rövid idejű memória sérülése – képtelenség a szótagtartásra

7. ábra

Összegző táblázat a kezelt betegekről és az eredményekről

Tizenkilenc afáziás beteget kezeltünk, akik neurológiai diagnózisuk és funkcionális vizsgálataink szerint globális afáziások voltak. A lejegyzett betegek között hét nő és tizenkettő férfi volt. Átlagos életkoruk 52,9 év, a fiatalabb és az idősebb öt betegen kívül a többség 50–60 év közötti volt. Az átlagos kezelési idő négy év, ami azt jelenti, hogy egy beteg eredményes kezeléséhez 420 terápiás órára volt szükség.

15 beteg kezelése eredményes volt. Ők bármely kommunikációs helyzetben szándékuk szerint, zavartalan információ-cserére képesek beszéd, olvasás és írás útján egyaránt.



4 beteg kezelése sikertelen volt. Ezeknél a betegeknél a nyelvi-, a pszichés- és a pszichoszociális adottságok bizonyultak kedvezőtleneknek. Eredménytelen-ségük több tanulsággal szolgált:

a) Az alacsony tanultsági szint és az alacsony szintű intellektuális befogadó képesség nem teszi lehetővé vállalkozásunk végrehajtását, amelyben a beteg nyelvi képességét kívánjuk újra felépíteni és/vagy előhívhatóvá tenni a betegben rejlő kognitív képességek függvényében.

Az egyik 70 éves analfabéta nő beteget hiába tanítottuk a hangok/betűk felismerésére, azok írására, ez mindvégig eredménytelen volt. Ugyanígy a globális szófelidézést sem támogatta a szóképes megerősítés. Végül a tárgyképek segítségével történő korlátozott számú szófelidézés sem adott lehetőséget a nyelvi képesség rendszerben működő helyreállítására.

Másik 70 éves nő betegünk négy elemi végzett, háztartásbeli volt és tanyájuk földműveléséből élt családjával. A beteg intellektuális befogadó képessége és intellektuális igénye hiányzott az olvasás- és írás alapú nyelvi terápiához.

b) A ritmus visszakopogásának a képtelensége két szempontból is kedvezőtlenül befolyásolhatta terápiánk sikerét. Pszichológiai szempontból a rövid távú akusztikus memória sérülésére következtethetünk, nyelvi szempontból pedig a belső ritmuson alapuló szótagolási képességet vélhetjük elveszítettnek.

Két 58 és 62 éves férfi beteg nem tudott ritmust visszakopogni, erre megtanítani nem tudtuk őket. Mindkét beteg beutalással, a sérülést követően másfél és két hónap után került terápiás foglalkozásra. Ekkorra kialakult hadaró, neologizmusokkal teli, érthetetlen beszédük. Készségük a verbális kontaktusra egyoldalú volt; nyelvi kifejezésre törekvő, a partner megértését mellőző. Tartós kontaktus kiépítésének hiányában terápiás próbálkozásaink az aktivizáló szakaszban eredménytelenek voltak. – A két eset hasonlósága arra is enged következtetni, hogy az arteria cerebri media elzáródás mindkét betegnél nagyobb vagy mélyebb területi sérülése a kezelhető betegekhez képest, makacs verbális és konstruktív apraxiát eredményezett. Erről a területi kiterjedésről érdemes lett volna computer tomographiás (CT) felvétellel meggyőződni. (A készülékkel a szegedi Röntgen Klinika csak 1990 első hónapjaitól rendelkezik.)

c) E két utóbbi esetben figyelemre méltó, hogy másfél, illetőleg két hónapig a betegek csak orvosi ellátásban részesültek és otthonukban családtagjaik értetlenségével találkoztak. A célirányos terápia kezdete késett. A terápia által javult betegek betegségük kezdetétől már a második, legkésőbb a negyedik héttől a célirányos terápia aktív részeseivé váltak. Véleményünk szerint a betegség kezdete

és a terápia kezdete között eltelt idő befolyásolja a beteg nyelvi felépülését (Darley 1982; Reinwang 1984; Lang – Stockert 1986).

### Irodalom

- Ádám György: A megismerés csapdái. Budapest 1987, 45–71.
- Darley, F. L.: Aphasia. Philadelphia, London, Toronto, Mexico City, Rio de Janeiro, Sydney, Tokio, 1982.
- Davis, G. A. – Wilcox, M. J.: Adult aphasia rehabilitation: applied pragmatics. Windsor, 1985.
- Cvetkova, L. S.: Some ways of optimisation of aphasics' rehabilitation. *International Journal of Rehabilitation Research* 3. 1980, 183–190.
- Finkelburg, F. C.: Asymbolie und Aphasie. *Berliner Klinisches Wochenschrift* 7, p. 1870, 449–450, 460–462.
- Froment, J.: La reeducation des aphasiques moteurs: principes, procedes et resultats. *Medical* 1921, 267–273.
- Fónagy Iván: Beszéd és valószínűség. *Magyar Nyelvőr*, 1986, 309–320.
- Geschwind, N. – Galaburda, M. A.: Cerebral lateralization. *Archives of Neurology* 42. 1985, 428–459.
- Glindemann, R. – Springer, L.: PACE-Therapie und Sprachsystematische Übungen. Ein integrativer Vorschlag zur Aphasietherapie. *Sprache-Stimme-Gehör* 13. 1989, 188–192.
- Glockman, J. M.: On increasing motivation to communication in aphasics' rehabilitation. *International Journal of Rehabilitation Research* 4. 1981, 78–81.
- Head, H.: Aphasia and kindred disorders of speech. Cambridge, 1926.
- Hegyí Ágnes: Neurolingvisztika a logopédiában. In: Művelődési Minisztérium Gyógypedagógiai Továbbképző Könyvtára. Budapest, 1988, 152–159.
- Helm-Estabrooks, N.A. – Fitzpatrick, P. M. – Barresi, B.: Response of an agrammatic patient to a syntax stimulation program for aphasia. *J. Speech Hearing Disord.* 46. 1981, 422–427.
- Howard, D. – Patterson, K. E. – Franklin, S. – Morton, J. – Orchard-Lisle, V. M.: Variability and consistency in picture naming by aphasic patients. In: F.C. Rose (ed.): *Recent Advances in Neurology*. 42. Progress in Aphasiology. New York, 1984.
- Howard, D. – Patterson, K. E. – Franklin, S. – Orchard-Lisle, V. M. – Morton, J.: The treatment of word retrieval deficits in aphasia: a comparison of two therapy methods. *Brain*. 108. 1985, 817–829.
- Kassai Ilona: A Magyar Fonetikai, Foniátriai és Logopédiai Társaság tájékoztatója 1. 1992, 5–9.
- Lang, Ch. – Stockert, Th. R.: Zum gegenwertigen Stand der Aphasietherapie. *Fortschr. Neurol. Psychiat.* 54. 1986, 119–137.
- Lebrun, Y.: Cerebral dominance for language. A neurolinguistic approach. *Folia Phoniatrica* 35. 1983, 13–39.
- Lytton, W. W. – Brust, J. C. M.: Direct dyslexia. Preserved oral reading of real words in Wernicke's aphasia. *Brain* 112. 1989, 583–594.
- Marshall, J. C. – Newcombe, F.: Patterns of paralexia. *Journal of Psycholinguistic Research* 2. 1973, 175–199.
- Merdian, G.: Semantische Störungen bei aphasikern. *Europäische Hochschulschriften*. P. Lang, Frankfurt am Main, New York, 1984.



- Newcombe, F. – Marshall, J. C.: On psycholinguistic classifications of the acquired dyslexias. *Bulletin of the Orton Society* 31. 1981, 29–46.
- Partz de, M. P.: Reducation of a deep dyslexic patient: rationale of the method and results. *Cognitive Neuropsychology* 3. 1986, 149–177.
- Pléh Csaba: Különböző szórendű mondatok értelmezése és a dichotikus hallási asszimetriák 3–6 éves gyermekeknél. *Pszichológia* 3. 1981, 365–393.
- Poeck, K.: Leitungsstörungen. In: K. Poeck (ed.) *Klinische Neuropsychologie*. Stuttgart, 1982, 159–167.
- Poizner, H. – Bellugi, U. – Klima, E. S.: Biological foundations of language. Clues from Sign Language. *Ann. Rev. Neurosci.* 13. 1990, 283–307.
- Reinwang, I.: The natural history of aphasia. In: Rose, F. C. (ed.) *Progress in Aphasiology. Advances in Neurology*, 42. New York, 1984, 13–24.
- Shewan, C. M. – Bandur, D. L.: Treatment of aphasia: a language-oriented approach. London, 1986.
- Weigl, I.: Neuropsychologische und psycholinguistische Grundlagen eines Programms zur Rehabilitierung aphasischer Störungen. In: G. Peuser (ed.) *Studien zur Sprachtherapie*. München, 1979, 491–514.

# AZ IDŐSKORI HANGKÉPZÉS JELLEMZŐI

**Balázs Boglárka**  
Szent János Kórház, Budapest

## Bevezetés

Az élettartam meghosszabbítása az emberiség régi vágyai közé tartozik. A betegségek megelőzése, a tünetek korai felismerése és az egyre nagyobb hatásfokú gyógykezelés évről-évre emeli az emberi életkor felső határát. Az ember öregedésével, öregkorának problémáival egyre több tudományág foglalkozik. A legrégebb és a legfejlettebb közöttük a gerontológia, az "öregségtudomány", amely minden lényeges biológiai, pszichológiai és társadalmi szempontból vizsgálja az időskort. Az öregedés ugyanis mindazon kémiai, fizikokémiai és morfológiai változásoknak az összessége és kölcsönhatása, amelyek az élő szervezet élete során átmegy és amelyek következtében élettani egyensúlyának állandó szintjét megtartani mind kevésbé képes. A gerontológia, mint az élettan egy része, eleinte az egész test morfológiai és funkcionális váltoásaival foglalkozott. Később áttért egyes szervek speciális öregedésének vizsgálatára. Legrégibb példája ennek a presbiopia – az öregkori látásromlás – és a presbiacusis – az öregkori hallásromlás kutatása. A WHO meghatározása szerint az öregkor 65 éves korral kezdődik. Hazánkban a 65 éven felüli lakosok aránya 18,5%.

Az öregedést Max Bürger határozta meg a legkorszerűbben: "Az öregedés biológiai folyamat. Az öregedés életfok, irreversibilis, amelynek oka és mechanizmusa van" (vö. Spellenberg 1988). A biomorfózis szakaszainak sorrendje megváltoztathatatlan és az idős szervezet megfiatalítása megoldatlan probléma. Ezért a gerontológiai kutatások és a geriátriai klinikum legfontosabb célkitűzése a minél hosszabb, ugyanakkor minél egészségesebb időskor elérése.

Hangképzőszervünk velünk együtt öregszik, ezért fontos, hogy ismerjük az egészséges öregkornak megfelelő hangképzési sajátosságokat és a hangminőség változásait.

## Foniátriai alapfogalmak

A hangképzés bonyolult folyamatának előfeltétele az idegrendszeri irányítás és ennek eredményeképpen a hangképző apparátus izmainak igen pontos koordinációja.

Az emberi hang keletkezésének alapja:

1. a kilégzés okozta levegőáramlás,
2. a hangszalagok fonációs beállítása, tónusa és a hangszalagok rezgése,



3. a toldalékcső rezonátori tevékenysége.

A hang minőségét konstitucionális, hormonális és pszihés tényezők befolyásolják.

Az emberi hang képzésének kérdése a mai napig sincs teljesen tisztázva. Két nézet terjedt el:

1. a régebbi, de ma inkább elismert aerodynamiás-muscularis elmélet és a
2. neurochronaxiás vagy neuromuscularis elmélet.

Az aerodynamiás-muscularis elmélet alapján a hangszalagok rezgése öngerjesztett rezgés. A fonációs állásban lévő hangszalagok megfeszülnek. A tüdőkből jövő levegő nyomása a hangrést periodikusan megnyitja. A hangszalagok megfelelő beállítása idegrendszeri hatás következménye.

A gége, vázizomzatának működésével, a hangszalagok rezgőképessége folytán és az egész gégeváz finom mozgásai által látja el hangképző munkáját. A belső gégeizmok mozgatják a hangszalagokat, illetve létrehozzák a hangszalagok fonációs beállítását. A tüdőkből kiáramló levegő a hangszalagokat rezgésbe hozza és hang keletkezik. A hangszalagok rezgőmozgása vízszintes síkú, de van egy kevésbé kifejezett függőleges komponens is.

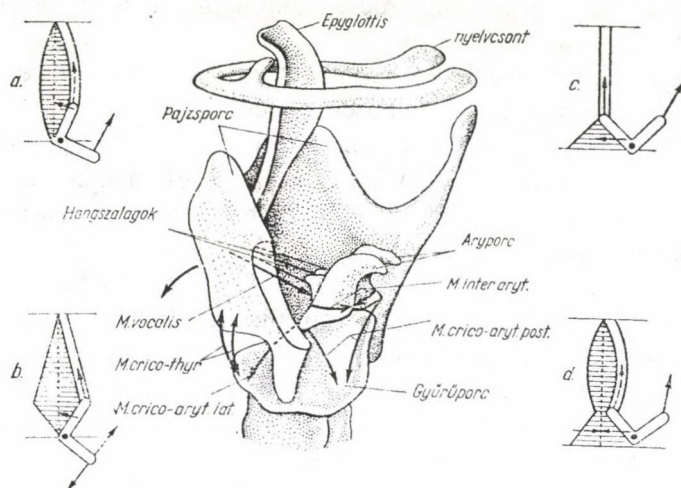
A hangszalagrezgés következtében kialakult hanghullám áterjed a gége alatti és feletti légtelt üregekbe és rezgéseit rákényszeríti azokra. A toldalékcső hangátalakításában az alábbi szerveknek van szerepük: a trachea és a bronchusok, a Morgagni-tasak, a garat, a nyelv, a légyszájpad, az orrüreg és az orrmelléküregek. A toldalékcső rezonátorláncként funkcionál, de falainak lágyrészei ugyanakkor egy kissé csillapítják is a hangot (1. ábra).

A **rekedtség** a hang tisztaságának zavara, akkor hallható, amikor a hangszalagrezgés által képződött tiszta zöngehanghoz zörejelemek társulnak. A zörejelemek a hangképzés folyamatának periodicitás-zavara miatt keletkeznek. A nem harmonikus rezgéskép kialakulásának két alapvető oka van: a hangszalagrezgés zavara és az áramló levegő turbulenciája.

A rekedtség vizsgálatát a szokásos fül-orr-gégészeti vizsgálaton kívül stroboszkópiával, glottográfiával és hangszinkép-clemezéssel: szonagráfiával végezzük. A hang dinamikájának vizsgálatára a fonetográfia szolgál.

### A hangképzés változása

A gége működése, egyben hangképzésünk, egész életünkben hormonális befolyásolás alatt áll. A megszületett gyermekek gégeje tekintet nélkül arra, hogy fiúk vagy lányok, egyforma nagyságú. A kisgyermekek hangmagassága, hangszíne, hangterjedelme és átlagos beszédhangfekvése lényegében egyforma. Ez a "gyermekhang".



1. ábra

A hangszalagok és a gégeváz ízületi mozgásainak térbeli ábrázolása

A változás a pubertás korában, a hormonális éréssel egyidőben történik: a mutációval elkülönül egymástól a női és a férfihang. A férfiak hangja egy oktávval mélyebb, férfias csengésű lesz, jellegzetesen eltérővé válik a női hangtól. Gerhard Böhme és Minor Gudrun vizsgálatai alapján ez az úgynevezett "felnőtthang" a pubertástól a 60. életévig hallható. A hatvanadik év után pedig évről-évre fokozatosan csökken a két nem hangminőségének távolsága – a férfiak és nők hangszíne, hangfekvése, valamint énekhangjuk milyensége – közeledik egymáshoz. A női hang magassága csökken, a jellegzetesen "nőies" hangszín kopik. A férfihang zengő érce halványodik, a férfias mélységek fokozatosan eltűnnek, a beszédhangmagasság emelkedik. Az idős emberek hangját "öreghangnak" nevezzük.

A presbiopia és a presbiacsis jelenléte az öregkorban jól ismert, a presbifónia szintén az öregkor velejárója. Ugyanakkor nem szabad figyelmen kívül hagynunk azt sem, hogy az életkor előrehaladtával az idős embereknek nemcsak a beszéd képzése, hanem a beszéd megértése is zavart szenved. Így vizsgálatainkban az egész szervezet valamennyi szervének kölcsönhatását kell szemügyre vennünk.



Az **idős emberek hangjának** változását mennyiségi és minőségi jellemzők adják. Az emberi hangképzést, a hangerőt, a hang tiszta csengését, színét, átlagos beszédhangmagasságát és az ember hangterjedelmét a hormonháztartás fiziológias és patológias folyamatai befolyásolják. A szervezet általános elhasználódása következtében, a természetes érlemeszedéssel, az izomzat működő tömegének csökkenésével, az idegrendszer fáradásával, valamint a csendesebbre állított, de összehangoltan működő belső elválasztású mirigyrendszer vezérlésével, lényegi különbségek tapasztalhatók az emberi hangban.

Mik ezek a változások?

A légzőrendszer öregedése következtében a tüdőkapacitás csökken, a szív munkája romlik. Tehát a gyengébben működő "pumpa" és az aerodinámias faktor hanyatlása miatt a **hangerő** csökken. Ehhez járul még hozzá a renyhébben dolgozó hangszalagzáró izomzat miatt csökkenő szubglottikus nyomás. Érthető módon ezek után a **hangtartás** is megrövidül. Ezért az idős ember hangja lényegesen halkabb, beszéd közben sűrűbben kell levegőt vennie. Az időskori hangképzésre a kisebb-nagyobb fokú erőtlenség a jellemző, amely miatt a hang reszketővé válik. A külső gégeizomzat leépülése, valamint a hangszalagokban végbemenő változások miatt a **hangmagasság** és a **hangterjedelem** is változik. A szövetekbe beépülő mészs megváltoztatja a stroboszkópos rezgésképet is. A hangszalagok rugalmasságának csökkenése az elasztikus rostok tönkremenetelét okozza. Ez férfiaknál a mély hangok területére esik: hangfekvéstől függően a F, G, A, H hangokra, míg nőknél a hangterjedelem felső részére vonatkozik a c2, d2, e2 hangokra. Így idős korban ismét a gyermekhangra jellemző bő 1 oktáv hangterjedelmet találunk, ellentétben a felnőtthangra jellemző, átlagosan 2 oktáv hangterjedelemmel szemben. A beszédhangmagasság a férfiaknál 1–1 1/2 hanggal emelkedik, míg nőknél vagy nem változik, vagy 1/2 hanggal mélyül.

Itt kell szót ejtenünk a dohányzás ártalmáról! A dohányzás miatt az évek alatt kialakuló kisebb-nagyobb megvastagodás, amely még nem tekinthető betegségnek, de a hangszalagok rezgő tömegét növeli, a hang mélyüléséhez vezet.

### Változások a toldalékcsőben

A hangszalagrezgés következtében kialakult hanghullám a toldalékcsőben alakul át az egyénre jellemző hanggá úgy, hogy saját rezgését rákényszeríti a hangszalagok alatt és felett elhelyezkedő rezonáns üregekre. A toldalékcsőben végbemenő változások tehát a hang minőségének változásához vezetnek.

A **trachea és a bronchusok** izomzatának és porcának meszesedése a mellkasi rezonancia mennyiségét és minőségét változtatja meg.

A **Morgagni-tasak** rugalmasságának csökkenése a hang akusztikai jellemzőit rontja.

A nehezkesebb nyelés és a **recessus piriformisok** részleges megnyílása a hangszínt fátyolosítja.

A **garat és gégeizomzat** fáradt működése részben a rezonanciát rontja, részben a hangszalagok fonációs állását gyengíti.

A **nyelv** mozgásának nehezítettége az artikulációt rontja. Ezenkívül a gége vázának alkalmazkodását – az emelkedést a magas hangok képzésénél és a süllyedést a mély hangoknál – is nehezíti a hangmagasság változtatásánál.

A **lágyszájpad** mozgása idős korban fokozatosan renyhül, így kisebb-nagyobb nyílt, sokszor kevert típusú orrhangzós színezet alakul ki az idős emberek beszédében.

Az **orrüregben** általában a nyálkahártya sorvadása figyelhető meg, amely szintén a rezonanciás viszonyok megváltozásához vezet.

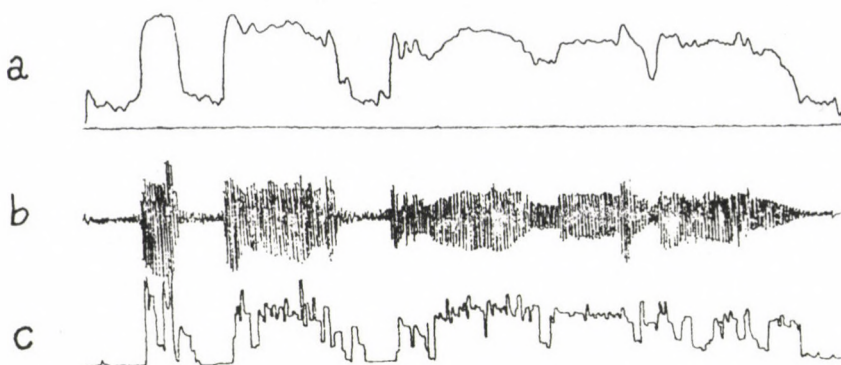
### A hangminőség változásai

A toldalékcsoében létrejövő, az öregkorban fiziológiásnak számító változások mind a hang tiszta csengését rontják. Akusztikai analízissel megállapítható, hogy a szonagramon a hangszínkép lényeges változása figyelhető meg. Sedlackova és a Gósy Máriával végzett saját vizsgálataink alapján az időskori hangra az a jellemző, hogy az alaphang kissé változik, és a szonagramon az élesen elhatárolt formánssávok helyett elmosódott szélű sávokat látunk. A harmonikusok száma csökken, a spektrum a magas frekvenciák irányába sokkal kevésbé terjed ki. A magánhangzók formáns-területén zörejelemek láthatók. Minél idősebb az ember, annál jellemzőbb a magas magánhangzó, főleg az [e] és az [i] harmonikus elemeinek felbomlása és a zörejelemek megjelenése. (Ennek okát a gerontológia a presbiacsis tényében látja, tehát a magas hangokon jelentkező halláscsökkenéssel, a minőségi hallásváltozással és a tökéletlen akusztikai visszacsatolással magyarázza.)

A glottogramon az intenzitás–idő függvényében az alaphangban olyan intenzitásváltozások vannak, amelyek azt mutatják, hogy a zöngék energiaviszonyaiban lökésszerű változások jelentkeznek. Jellegzetesen rövidülnek az amplitúdókiterések, és a rezgéskép a kifáradó izomzatra jellemző formát adja. (2. ábra).

A hang dinamikája is csökken. A fonetogram alakja, Wendler vizsgálatai alapján, érdekes módon a mutatóban látható alakváltozásra jellemző: a hangterjedelem és a hangerő csökken (3. ábra).





2. ábra

72 éves nő beszédhangjáról készült regisztrátumok  
(a=intenzitás, b=glottogram, c=alaphang)



3. ábra

Időskori beszéd fonetogramja (vonalkázott terület)  
egy átlagos, ép beszéd fonetogramjához viszonyítva  
(a szaggatott vonalak az átlagos beszédhangmagasságot jelölik)

Az a nagyon bonyolult folyamat, amely a beszédleégzésben, a hangszalagregzésben és a toldalékcső működésében hangadás alatt fennáll, egyrészt koordinált központi idegrendszeri, másrészt perifériás idegrendszeri és reflexes tevékenység következménye. Az életkor növekedésével az agyban lévő degeneratív elváltozások, másrészt a perifériás idegrendszer kopása megértési zavarokat is eredményez. A latencia idő megnövekedése a hangképzés és a beszéd minőségi romlását okozza. Az idős ember beszéde ezért kevésbé dinamikus, a beszéd színtelembb, monotonabb. Az emóciók kevésbé színezik a mondanivalót, bár ez egyes esetekben erősen személyiségfüggő. Magára a beszédre az egysíkúság és a pontatlan artikuláció jellemző. A beszéd állandóan elhalkul, feltöredezik. Ennek magyarázata Wendler szerint a szenzoros, extrapiramidális és limbikus-rendszer hangképzés irányításában történő csökkenő tendenciájában található.

Kísérletet végeztünk a beszédmegértéssel kapcsolatban, annál is inkább, mert hajlamosak vagyunk az idős emberek megértési nehézségeit aktuális hallásképességüknek, a presbiacusisnak betudni, illetve minden esetben hallászavart (is) feltételezni. Húsz, 70–80 év közötti idős emberrel rövid szövegrészt hallgattatunk meg egy meséből, majd öt kérdést tettünk fel nekik a hallott szöveggel kapcsolatban. A kérdéseinkre adott helyes válaszaik arányát az 1. táblázatban foglaltuk össze. Eredményeink azt mutatják, hogy a beszédmegértési nehézségek meglehetősen nagyok idős korban (természetesen adott életkortól, a tényleges hallásállapottól és intelligenciától, valamint a logikai készségtől is függenek a helyes válaszok). Ott, ahol a 25 éves átlagéletkorú kísérleti kontroll csoport 90–100%-os átlagteljesítményt nyújtott, ott a 70 év fölötti idősekből álló csoport átlagosan alig 51%-ot teljesített (1. táblázat).

1. táblázat

Időskori beszédmegértés	
Kérdés	Helyes válasz (%)
mikor?	16,6
miért?	91,6
hol?	87,5
mennyi?	33,3
mi az oka?	25,0

### Összehasonlító vizsgálatok

Tudjuk és tapasztaljuk is nap mint nap, hogy milyen hangzású az idős emberek beszéde. Ritkán vagy talán soha nincs azonban alkalmunk egymás után meghallgatni ugyanazon embernek a fiatal- és az időskori hangját.

Vizsgálatunkban módunk volt a Magyar Rádió segítségével meghallgatni és analizálni két színész fiatalkori és öregkori beszédhangját. Páger Antal hangfel-



vételei 1957-ben és 1985-ben, Dajka Margit felvételei pedig 1937-ben és 1983-ban készültek.

Nézzük meg, mit jelent akusztikailag a beszédben és a beszédhangban bekövetkezett változás?

A 2. táblázatban a legfontosabb akusztikai jellemzőket összegeztük a két színész beszédének alapján (f. a fiatalkori, ö. az öregkori beszéd elemzéséből átlagolt adatokat mutatja). A tempó változása szembetűnő: szavakban kifejezve ez azt jelenti, hogy Páger Antal fiatalkorában mintegy 40, Dajka Margit mintegy 20 szóval mondott többet egyetlen perc alatt.

Az alaphangmagasság a férfi ejtésben 15 Hz-cel emelkedett, a női ejtésben 160 Hz-cel csökkent időskorban.

Igen jellemző, hogy a túlsúlyban használt beszédhangok idős korban Páger Antalnál 500, Dajka Margitnál 1500 Hz-cel keskenyebb frekvenciasávban jelentek meg, mint fiatalkorban.

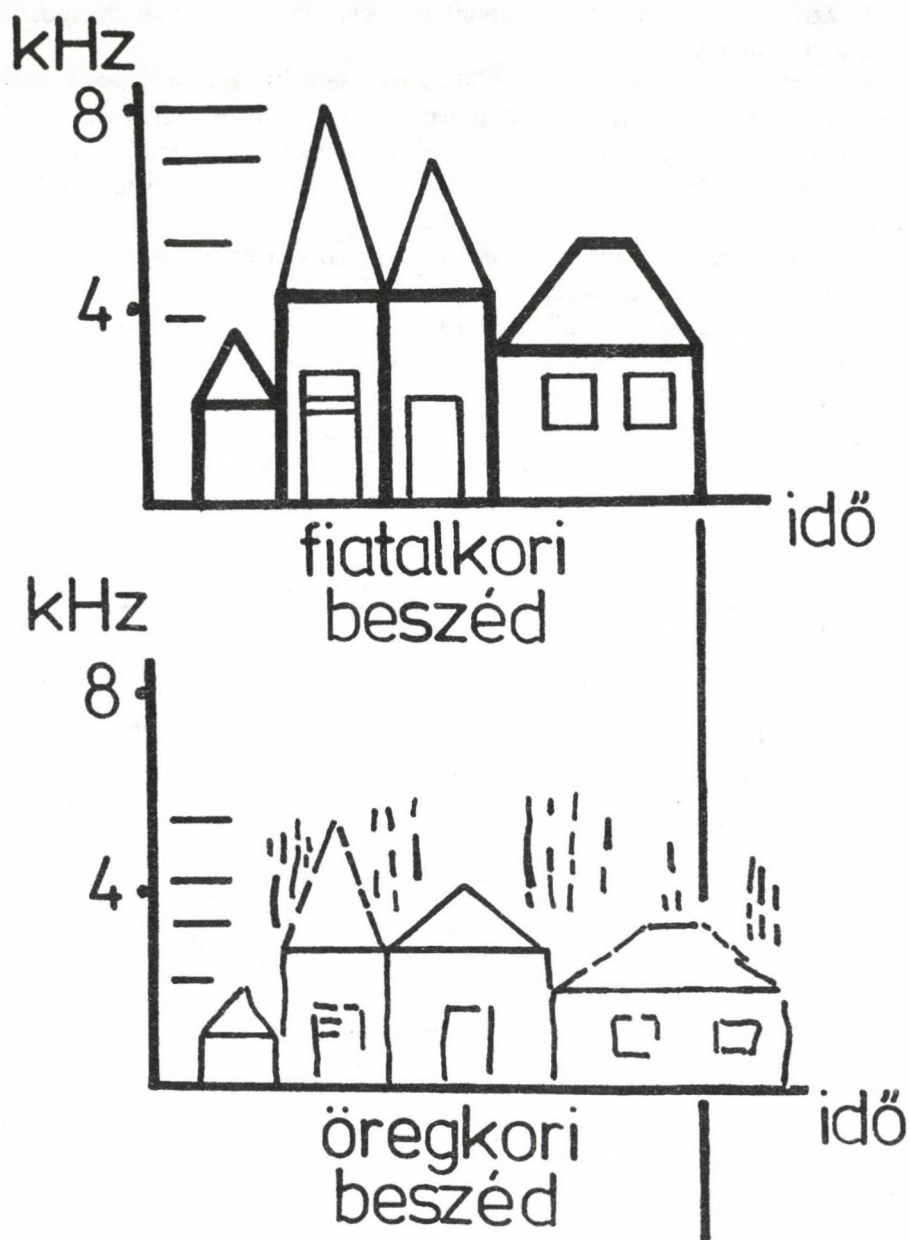
Igen jellegzetes a szünettartás alakulása: ha a férfi beszédének átlagos szünet-számát y-nak vesszük, látható, hogy időskorukban egyikük ötször, másikuk nyolcszor annyi szünetet használt beszéd közben.

2. táblázat

Vizsgált jellemzők	Kor	Páger	Dajka
tempó (h/s)	f.	12,3	13
	ö.	8,6	11
alaphang (Hz)	f.	90	280
	ö.	105	120
beszédfrekvencia (Hz)	f.	2500	3000
	ö.	2000	1500
szünettartás	f.	y	x
	ö.	8y	5x

Nyomdatechnikai nehézségek miatt az eredeti hangszíneképek helyett egyetlen sematikus ábrában igyekszünk szemléletesen összefoglalni az idős, illetve fiatalkori beszéd akusztikai sajátosságait (4.ábra).

A felső hangszíneképmítáció jelzi az erőteljes beszédhang komponenseket, a formánsok és a zörejösszetevők a mássalhangzók esetében kontúrosak, jól láthatók, megfelelően elkülönülnek egymástól, a teljes frekvenciasávban megjelennek. Ezek a fiatalkori beszéd akusztikai jellemzői. Az időskori beszéd akusztikumát reprezentáló alsó regisztrátum jelzi a formánsok intenzitásának csökkenését, a teljes beszédhullámon végigvonuló zörejelemeket, a bizonytalan formánsmegjelenést és az elmosódó kontúrokat.



4. ábra

A fiatal kori és az időskori beszéd akusztikumának sematikus ábrázolása



## Megbeszélés

A fentieket összefoglalva megállapíthatjuk, hogy a szervezet elhasználódásával együtt öregkorra nemcsak a hallás és a látás romlik meg, hanem a szépen csengő beszédhangunk is fátyolosabb, színtelenebb, reszkető és fáradékony lesz. Az élettartam meghosszabbításával ugyanakkor bizonyos emberi képességek nem csökkennek, sőt az egyén jó tulajdonságainak előtérbe kerülése biztosítja a tartalmas öregkort.

## Irodalom

- Balázs Boglárka: Hangképzési zavarok időskorban. Referátum. Miskolc. A Magyar Fonetikai, Foniátriai és Logopédiai Társaság kongresszusa, 1992.
- Balázs Boglárka – Gósy Mária: A beszéd vizsgálata gerontológiai szempontból. Debrecen. Előadás a Magyar Fül-Orr-Gége Orvosok nemzetközi kongresszusán, 1992.
- Balázs Boglárka – Gósy Mária: A beszédérthetőség és a beszédmegértés összehasonlítása idős korban. Nagykanizsa, Audiológiai vándorgyűlés, 1990.
- Böhme, G. – Hecker, G. (Minor, G.): Gerontologische Untersuchungen über Stimmumfang und Sprechstimmlage. *Folia Phoniatria* 22. 1976, 176–184.
- Childers, D. G. – Smith, A. G. – Moore, G. P.: Relationship between Electroglottograph, Speech and Vocal Cord Contact. *Folia Phoniatria* 36. 1984.
- Frint Tibor – Surján László: A hangképzés és zavarai, beszédzavarok. Budapest 1982.
- Kerényi Miklós György: Az éneklés művészete és pedagógiája. Budapest 1985.
- Nadoleczny, M.: Untersuchungen über den Kunstgesang. Berlin 1923.
- Schilling, A. – Karthaus, K.: Entwicklungsbeschleunigen und Stimmwechsel. *HNO* 9. 1960/61, 271–275.
- Schulz-Coulon, H.-J.: Die Diagnostik der gestörten Stimmfunction. *Archiv für Ohren-Nasen und Kehlkopfheilkunde* 227. 1980, 1–2.
- Sedlačkova, E.: Das Altern der Stimme. 7. Internat. Congr. Geront. Wien 1966. Congr. bericht Bd. 4, p. 469.
- Spellenberg, S.: A geroaudiológia és a gerootoneurológia aktuális kérdései. In: Az orvostudomány aktuális problémái. Budapest 1988.
- Wendler, J. – Seidner, W.: Lehrbuch der Phoniatrie. Leipzig 1987.

# **SZUBJEKTÍV ÉS OBJEKTÍV BESZÉDMINŐSÍTÉSI MÓDSZEREK**

**Molnár Sándor és Tatai Péter**  
Budapesti Műszaki Egyetem,  
Távközlési és Telematikai Tanszék

## **1. Bevezetés**

A hatékony beszéDMINŐSÍTÉSI módszerek szükségessége egyre fontosabbá válik a rohamosan szaporodó beszédkódolási rendszerek fejlesztése és alkalmazása céljából (BME-HEI). A számos módszer közül egy eredményes és gyors szubjektív minősítési módszert, az MNRU-t alkalmazó páros összehasonlításos módszert mutatjuk be, annak egy hatékony megvalósításával, a Tanszéken kifejlesztett Qualiphon rendszerrel. Az objektív módszerek közül a spektrum burkoló torzítási mértékek egy csoportját ismertetjük. A szubjektív és objektív vizsgálatokat elvégeztük egy egyre népszerűbbé váló, nemrég kifejlesztett beszédkódolón, a CELP (Code Excited Linear Prediction) kódoló 3200 bit/s-os és 4800 bit/s-os változatain. Az eredményeink alapján megkérdőjelezzük a széleskörben elfogadott kepsztrális torzítási mérték önmagában való eredményes alkalmazását torzítási mértékként, rámutatunk annak elégtelen voltára és javaslatot adunk az objektív beszéDMINŐSÉG kétparaméteres jellemzésére.

## **2. Szubjektív beszéDMINŐSÍTÉS**

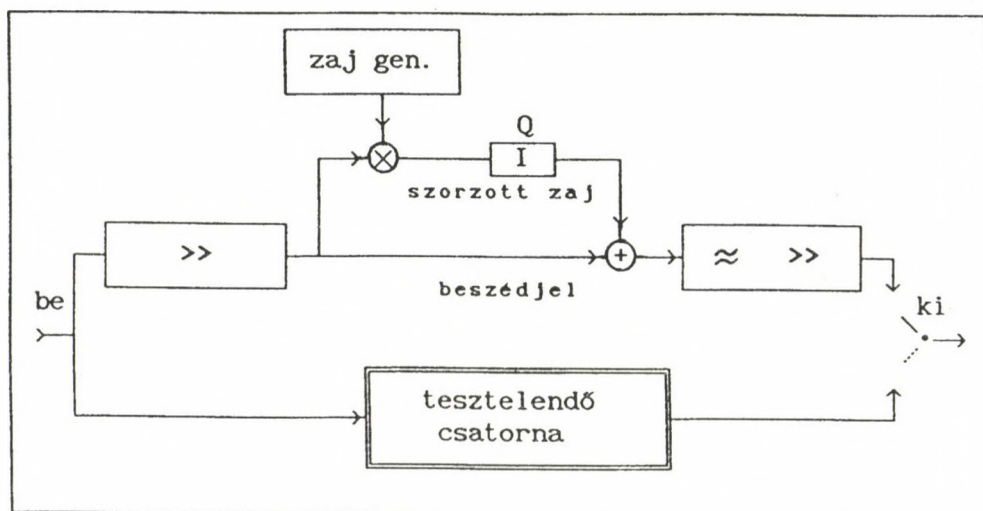
A szubjektív beszéDMINŐSÍTŐ módszerek olyan szabványosított eljárásokat használnak, melyek során a torzított beszédet emberi ítéletek alapján minősítik. A módszerek között két alapvető megközelítés figyelhető meg. Az egyik típus a beszéDMINŐSÉGET egyetlen számmal próbálja jellemezni, a másik megközelítés egy sokparaméteres jellemzést adja a beszédnek. Az előbbi előnye, hogy segítségével a kommunikációs rendszerek könnyen összehasonlíthatóak, de természetesen ezen módszerek nem adják a beszéDMINŐSÉG olyan pontos jellemzését, mint a sokparaméteres módszerek.

A beszéDMINŐSÉG meghatározására szolgáló módszerek között az egyik legelterjedtebb a Mean Opinion Score (MOS) (Kitawaki et al. 1984). Ez a módszer a beszéDMINŐSÉGET egy ötfokozatú skálával jellemzi (kiváló, jó, elfogadható, gyenge, rossz), mely nagyszámú hallgató ítéletének átlagaként alakul ki. A MOS egy nagyon jó és széleskörben elfogadott szubjektív beszéDMINŐSÍTÉSI módszer, de jelentős hátrányai is vannak: nagyon időigényes, nehézkes és drága módszer, továbbá gyengén torzított beszédminták között nem tud finom különbséget tenni.



## 2.1 A páros összehasonlításos módszer

A MOS-nál egy sok tekintetben ígéretesebb módszer a CCITT által javasolt modulált zaj referencia egységet (Modulated Noise Reference Unit, MNRU) használó páros összehasonlításos módszer (Noise Paired Comparison Test) (Kitawaki – Nagabuchi 1988). Az MNRU tartalmaz egy fehérzajforrást, melynek jelét összeszorozzák a beszédjellel, és az így előállított szorzott zajt kívánt szinten hozzáadják a beszédjelhez (1. ábra).



1. ábra

MNRU-t használó páros összehasonlításos módszer

Ez a referenciajel így tartalmaz egy beszéd komponenszt és egy beszédjellel modulált zaj komponenszt egyenletes frekvencia spektrummal. Egy csillapítóval beállítható a jel-zaj viszony ( $Q$ ), mely állandó a teljes dinamika tartományban. Az MNRU-t kiegészítve egy zajformáló szűrővel elérhetjük azt is, hogy a  $Q$  a frekvenciatartomány egyes sávjában is állandó legyen. A szokásos vizsgálati módszer szerint két mondatot játszanak le minden hallgatónak: egy torzított mondatot és egy referencia mondatot valamilyen  $Q$  értékkel. Minden hallgatónak azt kell eldöntenie, hogy melyiket tartja jobbnak a kettő közül. Az összehasonlítás "kényszerített", ez azt jelenti, hogy a hallgatóknak mindenképpen dönteniük kell, az "egyenlő minőség" válasz nem megengedett. A vizsgálatban a  $Q$  értékét változtatják egy tartományban. Az eredmények kiértékelésénél a torzított beszéd minőségét azzal a  $Q$  értékkel jellemzik, amelyhez az 50%-os preferencia szint tartozik.

A páros összehasonlításos módszer hátránya, hogy csak akkor alkalmazható eredményesen, ha a minőségromlás a vizsgált és a referencijelnél hasonló, mert az eltérő jellegű hatásokat nem lehet eredményesen összehasonlítani. Az MNRU torzítása a logaritmikus kvantálók torzításához hasonlít, mert  $Q$  a teljes dinamikatartományban független a jelszinttől; így például a logaritmikus kvantálók (szabványos PCM rendszerek) eredményesen vizsgálhatók ezzel a módszerrel.

A módszer előnye, hogy gyors és pontos jellemzését adja a beszédnek, már kis számú (kb. 20) hallgatóval végrehajtott tesztek is elég pontos eredményt adnak.

A vizsgálatnál a teszt előkészítése, a beszédanyag felvétele és az eredmények kiértékelése azok a feladatok, melyek sok időt igényelnek. Ezen szempontok figyelembevételével fejlesztettük ki a Qualiphon beszédminősítési rendszert (BME-HEI), mely a beszédminták összegyűjtését, a referenciamondatok előállítását és lejátszását a meghallgatási vizsgálat során, az eredmények kiértékelését stb. automatikusan elvégzi. A rendszer jelfeldolgozó processzorral megvalósított valós idejű MNRU-t tartalmazó IBM AT kompatibilis számítógépen működik. A szoftver Turbo Pascalban készült és többféle DSP-kártyát támogat.

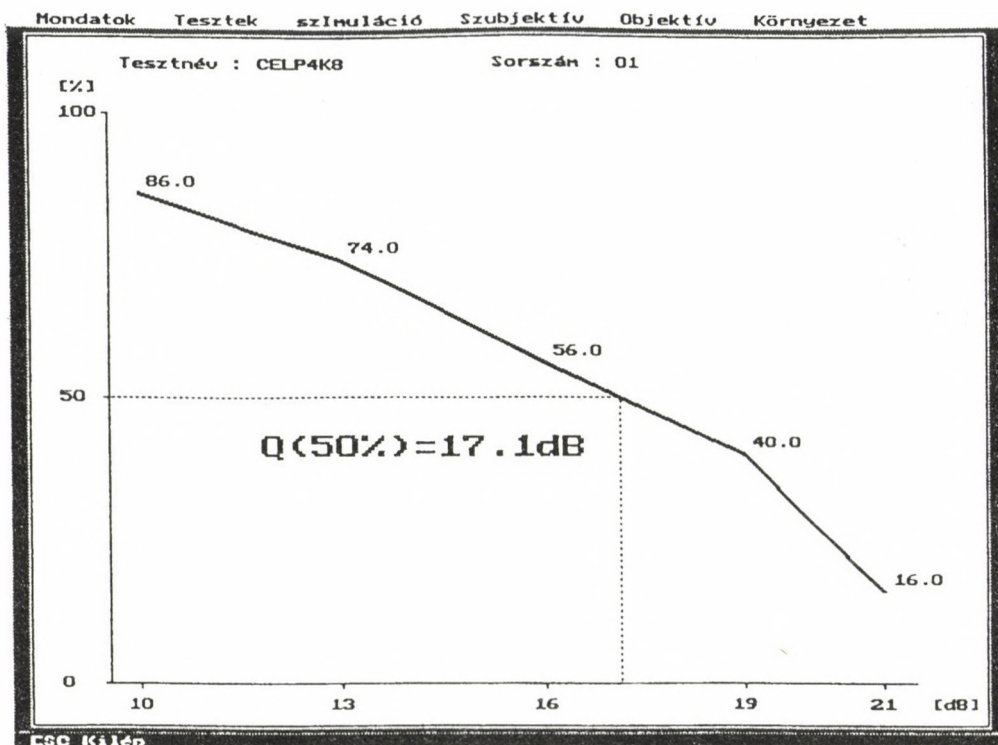
## **2.2 CELP kódoló szubjektív minősítése páros összehasonlításos módszerrel**

A vizsgálatainkban 3200 bit/s-os és 4800 bit/s-os CELP (Code Excited Linear Prediction) kódolót teszteltünk, melyek az amerikai DoD és AT&T Bell Laboratories által kifejlesztett Proposed US Federal Standard 1016 szabványnak felelnek meg (Hernández-Gómez et al. 1991).

A Qualiphon rendszerrel végrehajtott vizsgálatban (Molnár – Tatai – Jánosy 1992) 25 hallgató mindegyike kétszer 5 mondatpárból álló sorozatot hallgatott meg. A referencia és torzított beszéd sorrendje, valamint a referenciamondatok  $Q$  értékeinek sorrendje véletlenszerű volt. A mondatfüggőség elkerülése céljából a második sorozatban a referencia és torzított mondat generálási szabályát az első sorozat ellentétére változtattuk.

A preferencia kiértékelése a 4800 bit/s-os CELP kódoló esetében a 2. ábrán látható. Az eredmények átlagolása után a 3200 bit/s-os CELP kódolóra 11,7 dB, a 4800 bit/s-os kódolóra 17,1 dB adódott. A CCITT ajánlások szerint az elfogadhatóság alsó értéke 20 dB, mely 14 sorbakapcsolt 8-bites PCM kódoló minőségének felel meg (CCITT 1998). A vizsgált CELP kódolók így nem felelnek meg ezen követelményeknek, de az eredményekből várható, hogy a 7200 bit/s-os CELP kódoló valószínűleg már elfogadható eredményt fog adni.





2. ábra

A 4800 bit/s-os CELP kódoló preferencia tesztjének eredménye

### 3. Objektív beszédminősítés

Az objektív beszédminősítési módszerek az értékelési folyamatból elhagyják az emberi minőségítéletet (annak minden hátrányával együtt) és fizikailag mérhető paraméterekből számítanak beszédminősítési jellemzőt. Ebből következik, hogy az objektív beszédminősítési módszereknek számos előnye van a szubjektív módszerekkel szemben: gyorsak, egyszerűek, olcsók és ismételhetőek. Természetesen egy objektív módszer csak akkor használható, ha eredményei összhangban vannak a szubjektív eredményekkel. Egy ilyen objektív mérték konstruálása nem könnyű feladat, mert ehhez az emberi beszédérzékelési folyamatot

kellene jól ismernünk és modelleznünk. Ezen ismeretek hiánya miatt a jelenlegi objektív mértékeknek így a legnagyobb hátránya az, hogy eredményeik nem megfelelően korrelálnak a szubjektív eredményekkel.

Az objektív mértékek egy típusa a hullámforma torzítási mértékek, melyek az időtartományban vizsgálják az eredeti és torzított beszéd jelalakjának eltérését. Ezen mértékek a jel-zaj viszony valamilyen formáját alkalmazzák (hosszú idejű jel-zaj viszony, szegmentált jel-zaj viszony stb). A CELP kódoló esetében azonban nincs értelme ezen mértékek számításának, mert ez a kódoló nem hullámforma kódoló.

### 3.1 Spektrális burkoló torzítási mértékek

Másik nagy csoportját alkotják az objektív mértékeknek a spektrális burkoló torzítási mértékek. Ezen mértékek az eredeti és torzított beszéd spektrum burkolójának eltérését számítják. A beszéd spektrum burkolójának számítása lineáris predikció segítségével történik. A vizsgálatunkban a következő mértékeket használtuk:

#### a) Log Likelihood Ratio (LR)

Ez a mérték a torzítatlan és torzított beszéd csak pólusokat tartalmazó modelljeinek eltérését számítja a következő definíció alapján (Itoh et al. 1984):

$$LR = \left\{ 2 \ln \left( \frac{1}{\pi} \int_0^\pi \left| \frac{A_y(e^{j\omega})}{A_x(e^{j\omega})} \right|^2 d\omega \right) \right\}^{\frac{1}{2}} = \left\{ 2 \ln \left( \frac{\bar{a}_y R_x \bar{a}_y^T}{\bar{a}_x R_x \bar{a}_x^T} \right) \right\}^{\frac{1}{2}} \quad (1)$$

ahol  $A_x(z)$  és  $A_y(z)$  az eredeti és a torzított beszéd analízissz rőí, melyeket az  $\bar{a}_x$  és  $\bar{a}_y$  LPC egytthatók definiálnak.  $R_x$  az eredeti beszéd autokorrelációs mátrixa, melynek elemei a következők:

$$r_x(|i-j|) = \sum_{n=1}^{N-|i-j|} x(n)x(n+|i-j|) \quad \text{ahol } |i-j| = 0, 1, 2, \dots, p. \quad (2)$$

$N$  a minták száma egy szegmensben és  $p$  a predikció fokszáma.

#### b) LPC kepsztrum távolság mérték (CD)

A CD az eredeti és torzított beszéd kepsztrumtávolságán (Quackenbush et al. 1988) alapszik:



$$CD = \left( [c_x(0) - c_y(0)]^2 + 2 \sum_{k=1}^p [c_x(k) - c_y(k)]^2 \right)^{\frac{1}{2}} \quad (3)$$

ahol  $c_x(k)$  és  $c_y(k)$  az eredeti és torzított beszéd kepsztrumegei. A Parseval összefüggés alapján belátható, hogy a CD nem más, mint a kepsztrálisan simított spektrum logaritmusos spektrumtávolsága.

### c) COSH mérték (COSH)

Ez a mérték a torzítatlan és torzított beszéd közötti cosh függvény által súlyozott logaritmusos spektrumtávolságot adja (Itoh et al. 1984):

$$\text{COSH} = \left\{ \frac{2}{\pi} \int_0^{\pi} \left( \cosh[V(e^{j\omega})] - 1 \right) d\omega \right\}^{\frac{1}{2}} \Rightarrow$$

$$\min \text{COSH} = \left\{ 2 \left[ \frac{\sum_{i=-p}^p L_y(i) r_x(i) \sum_{i=-p}^p L_x(i) r_y(i)}{E_x E_y} \right]^{\frac{1}{2}} - 2 \right\}^{\frac{1}{2}} \quad (4)$$

$$\left[ \frac{G_x}{G_y} \right]^2$$

ahol

$$V(e^{j\omega}) = \ln \left( \frac{G_x^2}{|A_x(e^{j\omega})|^2} \right) - \ln \left( \frac{G_y^2}{|A_y(e^{j\omega})|^2} \right) \quad (5)$$

és  $L_x(i)$ ,  $L_y(i)$  a maximum likelihood együtthatók, melyek definíciói:

$$L_x(i) = \sum_{j=0}^{p-|i|} a_x(j) a_x(j+|i|) \quad (6); \quad L_y(i) = \sum_{j=0}^{p-|i|} a_y(j) a_y(j+|i|) \quad (7)$$

$E_x$  és  $E_y$  az eredeti és torzított beszédsegmens LPC maradéke energiái,  $G_x$  és  $G_y$  ezen szegmensek tiszta pólus modelljeinek erősítésfaktorai.

**d) Energy Ratio (ER)**

Az ER nagyon rokon mérték az LR-el (Quackenbush et al. 1988):

$$ER = \left( \frac{\bar{a}_y R_x \bar{a}_y^T}{\bar{a}_x R_x \bar{a}_x^T} \right)^{0,125} \quad (8)$$

**e) Lineáris és logaritmikus predikciós mértékek (LILPM, LOLPM)**

Ezek a mértékek a torzítatlan és torzított beszédsegmentumok LPC egytűtőinek lineáris és logaritmikus eltérését számítják (Quackenbush et al. 1988).

$$LILPM = \left[ \frac{1}{p} \sum_{i=1}^p [a_x(i) - a_y(i)]^2 \right]^{\frac{1}{2}} \quad (9)$$

$$LOLPM = \frac{1}{p} \sum_{i=1}^p 20 \log_{10} \left| \frac{a_y(i)}{a_x(i)} \right| \quad (10)$$

**f) Lineáris és logaritmikus reflexiós együttható mértékek (LIRCM, LORCM)**

Ugyanolyan típusok mint az előző mértékek, de ezek a PARCOR egytűtők lineáris és logaritmikus eltérését számítják (Quackenbush et al. 1988).

$$LIRCM = \frac{1}{p} \sum_{i=1}^p |k_x(i) - k_y(i)| \quad (11)$$

$$LOLPM = \frac{1}{p} \sum_{i=1}^p 20 \log_{10} \left| \frac{k_y(i)}{k_x(i)} \right| \quad (12)$$

**g) Lineáris és logaritmikus területarány (Area Ratio) mértékek (LIAR, LOAR)**

Számítási képletük ugyanaz, mint az előző mértékeké ((11) és (12)), de most  $ar_x(i)$  és  $ar_y(i)$  gynevezett Area Ratio egytűtőkkel, melyek definíciói:

$$ar_x(i) = \frac{1 + k_x(i)}{1 - k_x(i)} \quad (13); \quad ar_y(i) = \frac{1 + k_y(i)}{1 - k_y(i)} \quad (14)$$



Mindegyik mérték egy szegmensre számított LPC alapján ad eredményt. A teljes beszédre az egyes globális mértékek ezen mértékek átlagaként adódnak.

### 3.2 CELP kódoló objektív minősítése spektrum burkoló mértékekkel

A vizsgálatainkban 8 kHz-el mintavételezett beszédet használtunk. A lineáris predikciót 32 ms-os beszédsegmenthosszra alkalmaztuk Hamming ablakolással. Az eredmények az 1. táblázatban láthatók.

1. táblázat

	LR	CD	COSH	ER	LILPM
CELP 3k2	0,707685	1,741622	1,445977	1,097338	0,336789
CELP 4k8	0,727920	1,688726	1,463909	1,097704	0,349223

	LOLPM	LIRCM	LORCM	LIAR	LOAR
CELP 3k2	9,114059	0,102292	8,325775	0,342427	1,967899
CELP 4k8	9,485937	0,105472	8,596813	0,346012	2,020716

Az eredményekből látható, hogy mindegyik spektrális burkoló torzítási mérték azonos eredményt adott – a mérési és számítási hibán belül – mindkét kódolóra. Ennek magyarázata az, hogy ezek a torzítási mértékek a beszédben elsősorban a lineáris torzítást mutatják ki és az mindkét CELP kódolónál közel azonos.

## 4. Összefoglalás

A szubjektív és objektív vizsgálatok eredményeinek összevetéséből látható, hogy a szubjektív teszt által kimutatott jelentős minőségeltérést a két kódoló között a spektrális burkoló torzítási mértékek nem tudták kimutatni. A magyarázat az, hogy ezen mértékek a kisebb mérték nemlineáris torzítást, azaz például a torzított beszédben meglévő kvantálási zajt nem mutatják ki. Mindezek alapján nyilvánvaló, hogy önmagukban ezek a mértékek nem – még az egyre szélesebb körben elfogadott kepsztrális torzítási mérték (CD) sem – alkalmasak eredményes beszédminősítésre. Feltétlenül szükség van egy olyan objektív torzítási mértékre, amely az ilyen nemlineáris torzításokat képes kimutatni. Elgondolásunk szerint erre egy alkalmas módszer lehet az eredeti és torzított beszédjel analízissz rővel való kifehéritése utáni összehasonlítása, például szegmentált jel-zaj viszonytal. További vizsgálataink ennek a mértéknek a konstruálására irányulnak. Az eddigiek alapján nem várható, hogy olyan, a szubjektív eredményekkel jól összhangban lévő objektív mérték található, mely egyetlen paraméterrel jellemzi a beszédminőséget. Sokkal ígéretebbnek látszik a beszédminőség jellemzése két paraméterrel, melyek egyike a lineáris torzítást mutatja ki ha-

tékonyan (például a kepsztrális torzítási mérték) és a másik paraméter pedig a nemlineáris torzításra érzékeny.

### Irodalom

- BME-HEI: A digitalizált beszéd minősítéséről, OMFB Tanulmányok, 1987–1990.
- CCITT. Recommendations G. 101–G.181. Blue Book, vol. III.1 IXth Plenary Assembly, 1988.
- Hernández-Gómez, L. A. et al.: Real-Time Implementation and Evaluation of Variable Rate CELP Coders. IEEE Proc. Int. Conf. Acoust., Speech, Signal Processing, 1991, 585–588.
- Itoh, K. – Kitawaki, N. – Takehi, K.: Objective Quality Measures for Speech Waveform Coding Systems. Review of the Electrical Communication Laboratories, vol. 32. No. 2. 1984.
- Kitawaki, N. – Honda, M. – Itoh, K.: Speech Quality Assessment Methods for Speech-Coding Systems. IEEE Communications Magazine, vol. 22. no. 10. 1984, 26–33.
- Kitawaki, N. – Nagabuchi, H.: Quality Assessment of Speech Coding and Speech Synthesis Systems. IEEE Communications Magazine, vol. 26, no. 10. 1988, 36–44.
- Molnár, S. – Tatai, P. – Jánosy, Z.: Speech Quality Assessment for Low Bit-rate Coding. Journal on Communications, vol 43. 1992, 19–25.
- Quackenbush, S. R. – Barnwell, T. P. – Clements, M. A.: Objective Measures of Speech Quality. Prentice Hall 1988.
- Tatai, P.: Comments on Objective Quality Measures in Speech Encoding. Budavox Review 4. 1989, 20–24.



# TÉVESZTÉS ÉS MEGAKADÁS KÖTÖTT SZÖVEG FELOLVASÁSÁBAN

Molnár Ildikó

ELTE Bölcsészettudományi Kar, Budapest

## Bevezetés

1976-tól, szinte elindulásától (1973-ban volt az első verseny) figyelemmel kísérem a Pedagógusjelöltek Kazinczy-versenyét. 1987-től folyamatosan feldolgozom a versenyen elhangzó kötelező szöveg-interpretálásokat magnetofonfelvételek alapján választ keresve arra a kérdésre, milyen felolvasási - szövegmondási gondokkal küzdenek a leendő pedagógusok, illetve közülük azok, akik bizonyos válogatás után részt vettek – mint iskolájukban a legszebben beszélők – a Kazinczy-versenyen. Ez a munka, az 1992. évi szövegek feldolgozása a hatodik ilyen jellegű vizsgálatom (vö. Molnár 1988, 1991, 1992).

Amikor – közel két évtizede – a pedagógusjelöltek szép beszéd versenye elindult, a résztvevők legfőbb célja a szöveg pontos, mintaszerű felolvasása volt. A diákok igyekeztek szépen beszélni, szépen hangsúlyozott, szépen formált dallamú mondatokat mondani. Kevésbé sikerült a szöveg tartalmának, a gondolat egészének a megértése.

A versenyek második évtizedében, amikor a nyelvészeti kutatások a hangról, a szóról, a mondatról egyre inkább a szövegre kezdtek irányulni, a versenyzők teljesítményének értékelésében is előtérbe került a szövegmondás egészének összehatása, az értelemnek és a fülnek szóló együttes hatás, a szövegben rejlő gondolatok súlyviszonyainak érzékeltetése, a hangzásforma és a szövegforma összhangja. **Háttérbe szorult a tiszta artikuláció, a pontos olvasás igénye.** Az értékelés során is engedményeket tett a bírálóbizottság a szövegegész gondolati tartalmának érzékeltetése érdekében, és elnézte az apróbb pontatlanságokat. Ez az átcsúszás folyamatosan történt, de az utóbbi évekre állandósult a helyzet: kevésbé tiszta a versenyzők kiejtése és pontatlan a szövegolvasás (vö. Z. Szabó – Wacha 1978).

Magam először az 1988-as kötelező szöveg-tolmácsolások feldolgozása során figyeltem fel arra, hogy milyen kevés pedagógusjelölt olvas hibátlanul. Öt évvel ezelőtt a versenyzőknek már csak 30%-a olvasta pontosan a szöveget. 1992-ben is csak 29,7%-uk. Ez az arány tehát állandósulni látszik. Ezek az adatok elgondolkoztatóak. Hogyan nevelheti igényességre, pontosságra növendékeit a pontatlanul olvasó pedagógus?

Ez az észrevétel késztetett arra, hogy megvizsgáljam, milyen fonetikai jellegű botlások és tévesztések fordulnak elő a szövegfelolvasásokban 1992-ben.

## A vizsgálat anyaga

Az 1992-es kötelező szöveg Ady Endrének a Budapesti Naplóban 1906-ban megjelent publicisztikai stílusba tartozó írása volt, melynek címe: A hazaszeretet reformja (ld. Függelék). A vizsgálatot csak a folyamatos szövegre terjesztettem ki, a szerző neve és a cikk címe, valamint a lelőhely vizuálisan jól elkülönült, más betűtípussal, más nagyságban, más vastagságban szedték, s felolvasásukban nem is fordult elő tévesztés. A szöveg 320 szóból áll. 74 versenyző olvasta fel a szöveget (16 fiú, 58 leány), tehát a vizsgált korpusz 23 680 szóból áll. Az anyag 17 020 sec-nyi (4 óra 43 perc 40 másodperc) időtartamú.

A vizsgálat eredményeinek ismertetése előtt szólnom kell az olvasás körülményeiről is. A felolvasás célja általában a szövegben rejlő gondolatok értő és értető tolmácsolása. A Kazinczy-versenyen részt vevő felolvasónak azonban az elsődleges, az igazi célja az, hogy helyezést érjen el, érmet kapjon, dicsőséget szerezzen részint saját magának, részint iskolájának. Mindehhez csak eszköz a szöveg felolvasása. Nem mindennapi a szituáció, figyelemre méltóak az elhangzás körülményei is. A versenyző ünnepélyes külsőségek között olvas, szerepel. A hallgatóság a bírálóbizottság, a tanárok, a vetélytársak. A pedagógusjelölt sora kerülése előtt 15 perccel kapja meg a kötelező szöveget, így ennyi idő áll a rendelkezésére a felkészüléshez, a szöveg értelmezéséhez.

## Eredmények

1992-ben a kötelező szöveg felolvasása során 52 diák hibázott, rontott. A leányok közül többen olvastak pontosan, mint a fiúk közül (vö. 1. táblázat).

1. táblázat

	Felolvasók	
	Hibázók	Nem hibázók
fiú	13 (17,6%)	3 (4,1%)
leány	39 (52,7%)	19 (25,6%)
Összesen	52 (70,3%)	22 (29,7%)

A pontatlanul olvasók, hibázók közül 21 pedagógusjelölt csak egyszer rontott, de akadt olyan versenyző is, aki tizenegyszer tévesztett a szövegolvasás során (2. táblázat).

A feldolgozás során az észlelt hibákat (139 hiba) két nagy csoportra különítettem el, a **tévesztés** (I.) és a **megakadás** (II.) csoportjára. A vizsgált anyagban 97 tévesztés (az összes hibázás 69,8%-a) és 42 megakadás (az összes hibázás 30,2%-a) fordult elő.



2. táblázat

A hibák száma	A hibázók	
	száma	százalék
1	21	28,4
2	10	13,5
3	7	9,5
4	6	8,1
5	2	2,7
6	3	4,1
7	2	2,7
11	1	1,4
Összesen	139	100,0

I. **Tévesztés**, amikor a versenyző mást olvas és a félreolvasáson többé-kevésbé átsiklik. A tévesztés csoportjába tartozó hibák is tovább differenciálhatók. Mást olvas a felolvasó, de átsiklik ezen vagy azért, mert ő maga sem veszi észre, hogy nem az áll a papíron, amit olvasott, vagy észreveszi, de érvényesül a védekezési ösztön: abban bízik, hogy a zsüri nem veszi észre, s kevésbé tűnik ki a hiba, mintha javítaná. Hogy mikor melyik jelenségről van szó, az csak a hangzás, a magnetofonfelvétel alapján nem állapítható meg. Ha csak erre figyeltem volna a verseny folyamán, vagy jó videofelvétel állna rendelkezésemre, nonverbális jelzések talán segítenének az elkülönítésben, de az sem biztos.

Máskor a versenyző észreveszi, hogy hibázott és meg is ijed. Ilyenkor ez a megriadás megzökkentheti a felolvasás viszonylag egyenletes menetét. Jelezheti ezt az, hogy a versenyző kis szünetet tart az elrontott szó után, beleakad a nyelve a következő szóba, kirívóan megváltozik a beszélő hangszíne (elvékonyodik, fátyolos lesz a hangja stb.).

A rendszerezés során a tévesztés fajtáit 6 csoportba soroltam (3. táblázat).

3. táblázat

A tévesztések		
fajtája	száma	százaléka
hangtévesztés, hangcsere	2	2,1
kihagyás	12	12,4
beszúrás	12	12,4
a hasonló hangzású és hasonló jelentésű szavak cseréje	66	68
szóváltoztatás tartalmi, jelentésbeli hasonlóság alapján	4	4,1
szórendi változtatás	1	1

Nem soroltam a fenti típusok közé az idegen tulajdonnév (Jeanne d' Arc) betű szerinti vagy félig-meddig betű szerinti kiejtését. Mindegyik olvasatra akad példa! Ez nem felolvasási kérdés, a hibázó diák történelemismerete, műveltsége hiányos.

**1. A hangtévesztésnek 2 előfordulásával találkoztam.** Az egyik a tulajdonnév hibás olvasata: *Vázsolyival Vázsonyival* helyett. A másik az idegen szó félreolvasása, metatézis: *regilió* a *religió* helyén.

**2. Leggyakrabban a határozott névelőt hagyják el** a pedagógusjelöltek. Ez 1992-ben 9 esetben fordult elő. Elsősorban a gazdaságosságra törekvéssel magyarázható. A felolvasó úgy érzi, hogy a névelő elhangzása nélkül is érthető a szöveg, nem változik meg a jelentése. Néha tartalmasabb szavakat is kihagynak a versenyzők. Pl.: Pasteur nagyobb volt...Pasteur nagyobb *ember* volt helyett; e Párizsban... e *nagy* Párizsban helyett vagy Franciaországban ma másképpen szeretik a hazát... Franciaországban ma *már* másképpen szeretik a hazát helyett. Ezekben a mondatokban is érthető a szöveg, s nem sokat szegényedik a jelentés. Összesen 12 kihagyással találkoztam a feldolgozás során.

**3. Amilyen könnyen hagyják el, ugyanolyan nagyvonalúan is szúrják be** a határozott névelőt a szövegbe a felolvasók. A vizsgált anyagban 11-szer fordult elő ez. Az esetek nagyobb részében nem befolyásolják a félreolvasások a szöveg értelmét. Akad azonban ellenpélda is: Hogy az igazi evangélium szerint reformált hazaszeretet mi? A módosított változat bizony mást jelent: Hogy az igazi evangélium szerint *a* reformált hazaszeretet mi? Egy ízben előfordult más szó betoldása is: a ha tehetitek helyén ha *ti* tehetitek hangzott el. A beszúrások száma 12.

**4. A leggyakoribb tévesztéstípus az, amikor a szövegben szereplő szót hasonló hangzású és hasonló jelentésű alakváltozattal cseréli fel** a versenyző. Az e csoportba tartozó hibák száma 66. Általában akkor kerül sor a cserére, ha az alakváltozatok jelentése egészen közeli, pl.: *s* kötőszó helyén az *és* változat, *melynek* helyett az *amelynek*. 7 diák olvasta *leány* helyett a kevésbé választékos – s vélhetően a felolvasó által szívesebben használt – *lány* formát. A szövegben szerepel a *hazafiasság* és az *álhazafiság* szó. Ez nagyon megzavarja a pedagógusjelölteket. Mutatja ezt az is, hogy 10-szer fordul elő a *hazafiasság* helyén *hazafiság*, 3-szor pedig az *álhazafiság* helyén *álhazafiasság*. A *megrontott* religió helyén hallhatunk *megrontó* és kétszer *megromlott* religió változatot is. A boldogságra *termett* helyén 3 diák olvasott boldogságra *teremtett* szószerkezetet. Többször szerepel a Garibaldiak szent *hevüléséből* helyén a Garibaldiak szent *hevületéből* változat.

Sokszor csereberélik a jeleket és a ragokat is a felolvasók. Nyilván úgy vélik a pedagógusjelöltek, hogy nem sokat változtat az eredeti jelentésen, ha a többes



számú *iskolákban* alak helyén az egyes számú szó áll: *iskolában*. Ha jelen idejű igealak helyett múlt idejű szerepel: nem *tör* a házra senki helyett nem *tört* a házra senki, vagy ha beszűrnak egy birtokos személyjelet, így lesz a turpis *fejre* helyett turpis *fejére*. Előfordul más határozórag is: dolgozni az elnyomottak föl-szabadításában az elnyomottak föl-szabadításán helyett. Így is értelmes a szöveg: Hallottam egy párizsi liceum évzáró beszédéről egyszer, csak egy kicsit mást jelent, mint az eredeti: Hallottam egy párizsi liceum évzáró beszédében egyszer.

Baj van az igekötők pontos használatával is. Az eredeti *reformált* hazaszeretet jelzős szerkezet helyett 5 felolvasásban *megreformált* hazaszeretet szerepel, de van olyan versenyző is, aki a *formált* szóalakra cseréli az eredetit.

5. Előfordul az is, hogy **jelentésbeli hasonlóság alapján cserélik fel** a diákok a szavakat. Ilyenkor az egyes jelentések nagyon közel esnek egymáshoz: Franciaország mindig a népek élén *járt* helyett Franciaország mindig a népek élén *állt* hangzik 3 interpretálásban. Úgy tűnik, a felolvasó nagyobb valószínűséggel várja az általa választott szót, amikor a Tisza *táján* helyett a Tisza *partján* szerkezetet használja.

6. Ebbe a csoportba a **megváltoztatott szórend** esete került. Ilyenkor nagyjából azonos valószínűséggel várhatók a szavak, de a beszélő nyelvhasználatában inkább a megváltoztatott szórendű mondat fordul elő. Pl.: Nem nevelnek *itt ám* az iskolákban sem apró kozmopolitákat helyett nem nevelnek *ám itt...* szórend hallható. (A tévesztések egyes típusainak előfordulási arányát ld. a 3. táblázatban.)

II. A másik nagy hibacsoport a **megakadás**. A megakadások csoportját két alcsoportra különítettem: az **ismétlések** (42,9%) és a **téves kezdések** (57,1%) alcsoportjára.

Az **ismétlő** versenyző valamitől megijed, valami megakasztja, valami váratlan, szokatlan, ismeretlen hangalakú vagy jelentésű szó kerül a szeme elé. Ilyenkor a szöveg valamely részét még egyszer olvassa a diák. Az ismétlés általában rövid, leggyakrabban egy szótagnyi vagy annál alig hosszabb, s a kérdéses szóalakot bárhol ismételheti a felolvasó. Nem szükséges, hogy szótaghatáron vagy morfémahatáron álljon meg (vö. Maclay–Osgood 1977). A változatosságot mutatják az előforduló ismétlések: *el* – elpártolt, *elpá* – elpártolt, *elpár* – elpártolt, *megref* – megreformálták. Leggyakrabban szótaghatáron vagy morfémahatáron zökkenti meg a felolvasást a megakadó pedagógusjelölt: *köp* – köpködve, *cif* – cifrázva, *társá* – társadalom, *taní* – tanítók, *kozmo* – kozmopolitákat, *haza* – hazaszeretet, *szép* – szépséget. Ismételheti a megakadó az egész szót: *dolgozni* – dolgozni, *hazaszeretet* – hazaszeretet. Akad arra is példa, hogy az egész jelzős szerkezetet megismétli a pedagógusjelölt: *nemes társadalomnak* – nemes társadalomnak. Feltételezhető, hogy talán azért ismételi a versenyző, mert félreértelmezte, félreintonálta a mondatot. Ez a gyanú a feldolgozás során bennem is fel-

merült, ezért alaposan utána jártam, újra meghallgattam a szöveget, és megállapítottam, hogy nem erről van szó, mert az intonáció – füllel érzékelhetően – ugyanaz.

A megakadás másik alcsoportja a **téves kezdés**. Ilyenkor a felolvasó észreveszi, hogy hibázott, és a rosszul kiejtett szót, szótagot vagy szövegidegen szót újraolvasva, most már jól. A helyesbítés a Kazinczy-versenyek felolvasási gyakorlatában általában megjegyzés nélkül történik. (Mindössze egy esetben fordult elő, hogy a versenyző elnézést kért, amint észrevette a hibát. Egy diák pedig azt kérte, hogy hadd kezdje újra a szövegolvasást.)

A feldolgozás során a téves kezdésnek, a szövegidegen nyelvi elem korrigálásának 5 típusát különítettem el (4. táblázat).

4. táblázat

A tévesztések		
fajtája	száma	százaléka
hangtévesztés	5	20,8
hasonló hangalakú szavak cseréje	4	16,7
téves kezdés hasonló hangzású és hasonló jelentésű szavak esetén	11	45,8
téves kezdés hasonló jelentésű szavak esetén	2	8,3
szórendi változtatás	2	8,3

**1. Hangtévesztés** 5 esetben fordult elő. Ilyenkor javíthat a felolvasó mindjárt a tévesztett hang után megállva: *cip* – cifrázva, *álhas* – álhazafiasság, vagy az egész hangalak kiejtése után, így: *rabró* – rabló.

**2. A hasonló hangalakú szavak** úgy kerülhetnek bele a szövegbe, hogy felületesen rápillantva a nyomtatásra, nem figyel a jelentésre a felolvasó, s csak a következő szóhoz érve veszi észre, hogy félreolvasott. Így lesz altató zsoltára helyett *apró* – altató zsoltára, máskor pedig: Nem a régi: az *üreg* – az üres, a hazug, a sallangos. Ebbe a csoportba 4 esetet soroltam.

**3. A téves kezdések** legtöbb előfordulása ebbe a csoportba tartozik. 11 példát gyűjtöttem ki az anyagból. Nem kell itt sem az egész rontott szót elolvasnia a versenyzőnek. Jól példázzák ezt a következő előfordulások, és megmutatják azt is, hogy melyik változat felelt volna meg a felolvasó előzetes várakozásának. Nyilván *hevületéből* szót várt a pedagógusjelölt, amikor így rontott: *hevüle* – hevüléséből. Egy másik diák hadvezér szóra számított, amikor ezt olvasta: *hadvez* – hadverő hős. A Garibaldiak analógiájára bizonyára csak családi nevet várt a



felolvasó, amikor így olvasott: mi lesz a *Kossutho* – Kossuth Lajosok és a Garibaldiak szent hevüléséből.

Ennek a típusnak alfaja a képző-, rag-, illetve jeltévesztés is: boldogságra *terremett* – termett minden ember vagy *Vázsonyinál* – Vázsonyival.

4. Ebbe a csoportba azok az esetek tartoznak, amikor a **hasonló jelentésű** szókapcsolatok közül a gyakoribb formát szeretné használni a felolvasó. Franciaországban ma már másképpen szeretik a hazát, mint például a Tisza *pa* – táján. Ez a hiba 2 esetben fordult elő.

5. A típus hibázói saját gondolatmenetüknek, mondat szerkesztésüknek megfelelő **szórendet** várnak. Az egyénibb szórend megzavarja őket. Pl.: Csak egy kicsit megreformálták azt a megrontott religiót, melynek *hazafiasság* a – melynek neve hazafiasság. Ilyen szórendi rontásra 2 példát találtam.

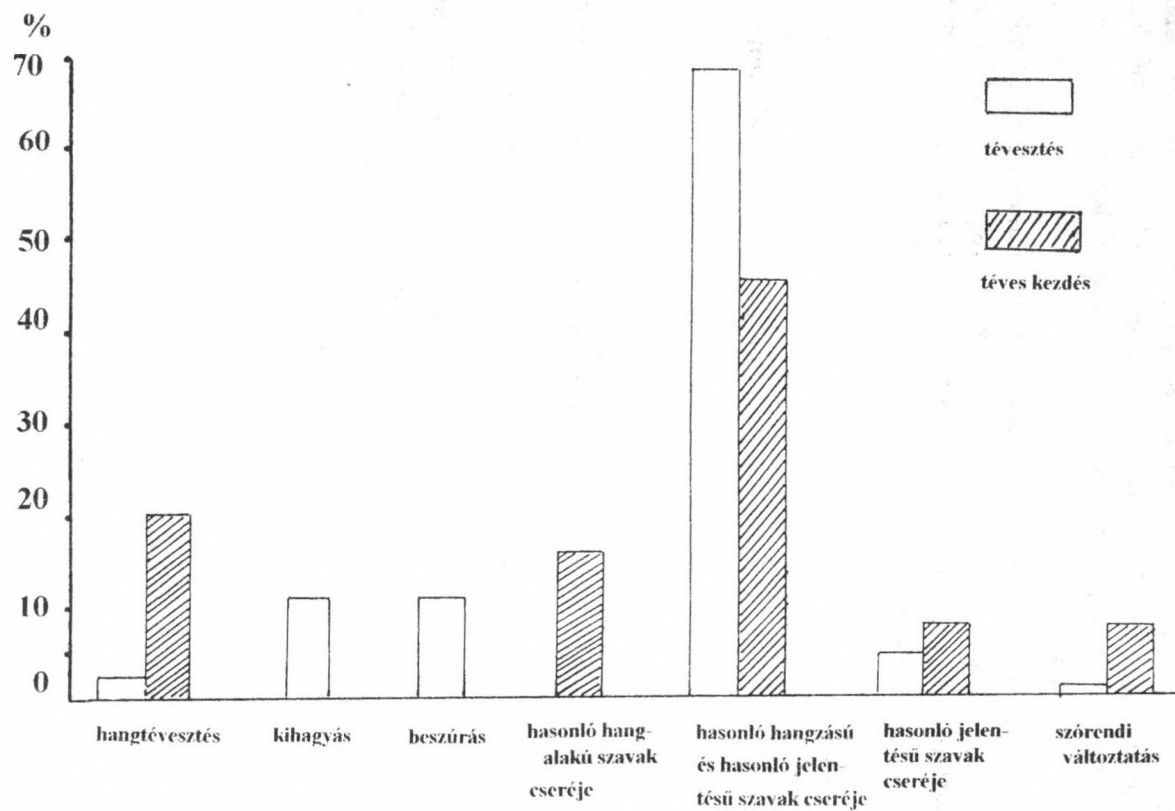
### Tanulságok

A hibák okait kutatva érdemes megnézni, hogy vannak-e a szövegben olyan helyek, ahol sűrűsödnek a tévesztések és a megakadások. Megállapítható, hogy vannak. Megzavarja a felolvasókat:

- a) az idegen szó (kozmpolitákat, megreformálták, religiót),
- b) az ismeretlen jelentésű szó (turpis),
- c) az ismeretlen tulajdonnév (Vázsonyi),
- d) az olyan idegen tulajdonnév, amelyet csak felületesen ismer a versenyző és kiejtésében bizonytalan (Jeanne d' Arc).

Valószínűleg ezeknek a szavaknak a kiolvasásától, kiejtésétől már előre fél a verseny résztvevője.

Végezetül verbális összefoglalás helyett álljon itt egy diagram (1. ábra) az 1992-es Kazinczy-verseny kötelező szöveg-interpretációiban hallható fonetikai jellegű tévesztésekről és megakadásokról.



1. ábra  
Fonetikai jellegű tévesztések és megakadások



## Függelék

Ady Endre:

### A HAZASZERETET REFORMJA

Ez Úristentől elpártolt Franciaországban nem tör a Hazára senki. ...

Nem nevelnek itt ám az iskolákban sem apró kozmopolitákat. Csak egy kicsit megreformálták azt a megrontott religiót, melynek neve hazafiasság. Franciaországban ma már másképpen szeretik a hazát, mint például a Tisza táján. ...

Franciaországban hamar észrevették, mi lesz a Kossuth Lajosok és a Garibaldiak szent hevüléséből. Rabszolgatartók altató zsoltára, fekete seregek csillogó vértje, hitvány törtetők létrája s gazemberek menedéke, mint főtisztelendő Szerényi úr itthon a napokban néhány turpis fejre ráidézte.

Nem éltetve, dalolva, cifrázva, köpdösve s részegen szeretik itt a hazát. És nem a múltját szeretik, hanem a jövőjét. A dicsőséges XIV. Lajosról azt tanulják a gyerekek, hogy szívtelen rabló volt. Jeanne d' Arc egy eszelős, hisztériás leány, s Napóleon egy új Attila. Pasteur nagyobb ember volt, mint tíz hadverő hős. Ilyeneket tanítanak bölcsen a francia nebulóknak, s nálunk nemrég kenyérmezei táncot jártak Vázsonyival Kinizsi Pál miatt.

Hogy az igazi evangélium szerint reformált hazaszeretet mi? Hallottam egy párizsi líceum évváró beszédjében egyszer:

– Fiúk, nézzetek szét e nagy Párizsban s az egész világon, ha tehetitek. Szívetek és szemetek meg fog telni könnyel. Mennyi galádság, tudatlanság, inség és bűn. Mennyi még az elnyomottak száma, mikor pedig boldogságra temett minden ember. Dolgozni az elnyomottak fölsszabadításán, minden ember boldogságán: ez a cél. Agyak és kezek kooperatív, szent munkája: ez az eszköz. Igazságot, világosságot, boldogságot és szépséget adjunk a Föld minden emberlakójának. Meg kell csinálnunk az igazságos, új és fölséges emberi társadalmat. Fiúk, legyetek büszkék, hogy franciák vagytok, Franciaország mindig a népek élén járt az emberiség boldogításáért. Az új, a szabad, a nemes társadalomnak mintáját a világ számára Franciaországnak kell megteremtenie.

Íme, az új hazafiasság: az emberi, a populáris, a szociális hazafiasság. A XX. század emberéhez illő. Nem a régi: az üres, a hazug, a sallangos. És Franciaország legújabb generációja már e vallásban nevelkedik.

... És nincs más útja a társadalmi haladásnak, csak ez: új tanítók és új iskolák. Mindenekfölött pedig: a hazaszeretet reformja s szent háború az álhazafiság ellen.

### Irodalom

- Maclay, H. – Osgood, Ch.: Hezitációs jelenségek a spontán angol beszédben. In: Pszicholingvisztika és kommunikációkutatás. Budapest 1977, 97–132.
- Molnár Ildikó: Hezitációs jelenségek az élőbeszédben. MFF 3. 1979, 49–57.
- Molnár Ildikó: A pedagógusjelöltek beszédének hangzástani jellemzői az 1988. évi Kazinczy-verseny hangfelvételeinek elemzése alapján. EFF 2. 1988, 140–6.
- Molnár Ildikó: A szegedi országos döntő kötelező szövegének hibaelemzése. EFF 4. 1991, 125–31.
- Molnár Ildikó: Szöveg és hangzás együttese 1991-ben Sárospatakon. EFF 5. 1992, 47–50.
- Szende Tamás: Megérthetjük-e egymást? Budapest 1987.
- Z. Szabó László – Wacha Imre: A Kazinczy-versenyek története. Győr 1978.



# A TEMPÓ ÉS A SZÜNET VISZONYA A HANGOS OLVASÁSBAN

**Laczkó Mária**

MTA Nyelvtudományi Intézete, Budapest

A beszéd folyamat Goldman-Eisler (1968) felfogásában a hangsorok és a szünetek folyamatos egymásutánisága, ritmikus váltakozása. Maclay – Osgood (1959) azt is megállapítja, hogy a folyamatot az artikuláció elmaradásakor nemcsak a hangos elemeket nem tartalmazó szünetek, hanem akusztikailag azonosítható hangjelenségek is kitölthetik.

Az írott szöveg felolvasását – az ún. interpretatív közlést – szintén megszakításokkal tagolják. Ennek a tagolási funkciónak a célja a szöveg gondolati-logikai-nyelvi viszonyainak tükröztetése, amelyet a felolvasó a mondatfonetikai eszközök (a hangerő, a hangmagasság, a dallam, a tempó váltása, a hangsúlyozás formái) változtatásával valósít meg. A szöveg tagolásában fontos szerepet töltenek be a szünetek is, amelyek az olvasáshoz szükséges alap- és pótlevegőt biztosítják.

A szünetek típusainak spontán beszédbeli megjelenését a beszédhez szükséges levegő pótlása mellett gyakran hozzák kapcsolatba a gondolkodás és a beszédtevékenység összefüggései alapján a beszédtervezési műveletekkel is (Maclay – Osgood 1959; Boomer 1965; Butterworth 1989).

Az írott szöveg megelevenítésekor a mondanivaló pontosan ismert a beszélő számára s így kevesebb tervezési műveletre van szükség, ezért feltehetően kevesebb a szünetek száma is. Számos tanulmány rámutat arra, hogy a szünetek az interpretatív közlésben nemcsak ritkábban, hanem meghatározott helyeken jelentkeznek, ott, ahol azt a mondat szerkezete lehetővé teszi, s szinte kizárólag légzési szünetként funkcionálnak (Maclay – Osgood 1959; Goldman-Eisler 1968; Martin 1970). A szünetek száma és időtartama az artikulációs sebességgel együtt meghatározó a beszéd, illetőleg a felolvasás tempójának kialakításában (vö. Fónagy – Magdics 1960; Gósy 1989; Váradi 1988; Gósy 1991).

A felolvasás közbeni szünetek és a tempó megítélésekor más jellemzőkkel is számolnunk kell. Az írott szöveg grafikailag bekezdésekre tagolódik, s ezeket, valamint az őket felépítő grammatikai, szintaktikai egységeket (mondatok, tagmondatok, szószerkezetek, szavak) a szóközön kívül az írásjelek és/vagy kötőszók választják el és kapcsolják össze. Ezek nyelvtani, grammatikai jelzések, nem mondatfonetikai eszközök, következésképpen az értelmi tagolásra való törekvés miatt, nem törvényszerű minden esetben szünetként való értelmezésük. Vizsgálatok azonban igazolták, hogy a felolvasók hajlamosak a szünetek eloszlá-

sát a vizuális megerősítéshez igazítani, s így a mondattani és intonációs egységek alapján várható és megjósolható szünetek eloszlása megbomlik (vö. Goldman-Eisler 1968). A tagmondatok közötti szünetek valószínűségét továbbá a szintaktikai szerkezet is befolyásolja (Hegedűs 1953; Fónagy 1967), a tempóval kapcsolatos kutatások pedig a tempó gyorsulását igazolták a magyar nyelvben (Gósy 1991; Kassai ebben a kötetben).

Jelen tanulmányban a felnőttek hangos olvasása közben tartott szüneteket elemeztük. Vizsgálatunk céljai a következők voltak: 1. a felolvasás tempója, valamint a tempó változtatása hogyan jellemezhető az artikulációs tempóval és a szünetek arányával 2. milyen a szintaktikai szünetek előfordulása a gyakorlott és a kevésbé gyakorlott beszélők szövegolvasásakor 3. módosul-e és hogyan az ún. grammatikus szünetek (vö. Fónagy 1967) megjelenése a kötőszóknak, a központosítás vizuális hatásának és a mondat szerkezetnek a függvényében és ez hogyan jelentkezik a tempó változtatásakor.

### **Anyag és módszer**

Három kísérletsorozatot végeztünk el. A kísérlet anyagául választott szöveg minden esetben megegyezett, s természetesen a magyar nyelvre jellemzően alárendelő és mellérendelő mondatokat és szó szerkezeteket, valamint kötőszóval és anélkül kapcsolódó tagmondatokat, szó szerkezeteket egyaránt tartalmazott. Az első és a második kísérletben a szöveg úgy volt nyomtatva, hogy a központosítás alapján bekezdésekre, mondatokra, tagmondatokra tagolódott. Ebben a két kísérletben 5-5 gyakorlott (I. csoport) és átlagos beszélő (II. csoport) vett részt. Feladatuk a szöveg hangos felolvasása volt először az egyénre jellemző természetes tempóban, majd ehhez képest gyors, végül pedig lassú felolvasást kértünk. A harmadik kísérletben 5 átlagos beszélő (III. csoport) szerepelt, nekik először a központosítást kellett jelölniük, majd a szövegnek ismét háromféle tempójú hangos felolvasását kértük. A kísérletsorozatban részt vevő 15 fő minden csoportban azonos nem szerinti bontásban szerepelt, átlagéletkoruk 38 év volt. Minden adatközlő hangos szövegolvasását magnetofonra rögzítettük, majd a felvett anyagról minden esetben intenzitásmérővel intenzitásgörbékét készítettünk. A regisztrátumokkal és a szöveg lehallgatásával párhuzamosan elemeztük a felolvasás közbeni szüneteket.

### **Eredmények, következtetések**

Elsőként arra kerestünk választ, hogy az olvasás tempója és annak gyorsítása vagy lassítása az 1 s alatt kiejtett beszédhangok számával és/vagy a szünetek számával és időtartamával jellemezhető-e inkább. Az 1. táblázat a beszédtempó (BT) és az artikulációs tempó adatokat (AT) mutatja mindhárom olvasási sebesség esetén. (A kétféle tempó elkülönítésére vö. Gósy 1991.) A BT számításakor



a szakirodalomban általánosan elfogadott definíció (vö. Maclay – Osgood 1959; Goldman-Eisler 1968; Sallai – Szende 1975; 1979; Váradi 1988) szerinti csendes szüneteket vettük figyelembe. Ennek alapján szünetnek tekintettünk minden olyan jelkimaradást, amelyik legalább 100 ms hosszúságú volt. (A meghatározás szerinti hangos szünetek aránya anyagunkban nagyon csekély volt.)

Az átlagos beszélőknél az a tény, hogy a szöveget tagolniuk kellett, nem mutatott lényeges eltéréseket ugyanezen beszélők másik csoportjához képest, így a tempóadatokat a gyakorlott és az átlagos beszélők szerinti bontásban tüntettük fel.

1. táblázat

Az olvasás tempója	Beszédtempó		Artikulációs tempó	
	hang/s			
	gyakorlott	átlagos	gyakorlott	átlagos
	beszélők			
lassú	8,99	9,36	12,12	11,88
természetes	11,22	10,95	14,45	13,42
gyors	15,40	13,98	17,31	15,67

1. A gyakorlott és az átlagos beszélők tempóértékei valamennyi ejtésben nagyjából megegyeztek. (A két csoport között a gyors ejtésben mutatkoztak a legnagyobb különbségek a BT és az AT esetén is, de ezek is csak 1,4, ill. 1,6 hang/s nagyságrendűek voltak.)

2. A BT és AT közötti különbségek sem mutattak lényeges eltérést a két csoportban. A gyakorlott beszélőknél azonos (megközelítőleg 3 hang/s) volt a különbség a lassú és a természetes ejtésben, a gyors ejtésben ehhez képest valamivel kisebb, alig 2 hang/s. Az átlagos beszélőknél a kétféle tempóérték között kisebb különbségekkel ugyan (kb. 2,5 hang/s volt a különbség a lassú és a természetes ejtésben, s közel 1,5 hang/s a gyors ejtésben), de hasonló tendencia volt kimutatható.

3. A természetes tempóhoz viszonyítva a lassításkor kisebb, a gyorsításkor pedig nagyobb különbségek mutatkoztak a BT-ben és az AT-ben is. A különbségek jelentősebbek a BT-ben, (lassításkor 2,23 hang/s, gyorsításkor 4,1 hang/s a gyakorlott beszélőknél, az átlagos beszélőknél pedig közel 1,6 hang/s, illetőleg 3,03 hang/s), a legnagyobb különbség szintén a BT-ben volt gyors tempóban (4,1 hang/s). A lassításkor és a gyorsításkor mutatkozó különbségek a két csoportban ezúttal is hasonlóak voltak.

Összegezve, s összehasonlítva ezeket az adatokat a szakirodalomban található olvasási tempóértékekkel (vö. Fónagy – Magdics 1960, 185) feltételezhető, hogy a magasabb olvasási tempóértékek a felgyorsult beszédtempónak az olvasásra gyakorolt hatásával is magyarázhatók. Másfelől kimondható, hogy az olvasás

tempójában nem érvényesül az adatközlőknek a beszédben vagy az írott szövegben való jártassága. Az olvasási sebesség sokkal inkább jellemezhető a felolvasás közbeni szünetek időtartamával, mint az 1 s alatt kiejtett beszédhangok számával. A tempó változtatásakor a gyorsítás főleg a szünetek időtartamának csökkenésében, a lassítás pedig a szünetek számának növekedésében jelentkezik.

A szöveg olvasására fordított időnek és a szünetek időtartamának egymáshoz viszonyított százalékos aránya a 2. táblázatban látható. Ahogyan az feltételezhető volt, mindkét időtartam a természetes tempóhoz viszonyítva a tempó változtatásával egyenes arányban változott. A százalékos adatok a gyakorlott és az átlagos beszélőknél nagyjából megegyeztek valamennyi tempóban, a természetes tempóhoz képest pedig mind a két csoportban a gyors tempóban mutatkoztak jelentős különbségek.

2. táblázat

Az olvasás tempója	A szünet aránya a beszédben	
	gyakorlott beszélők	átlagos beszélők
lassú	25,10	20,77
természetes	22,15	18,77
gyors	11,12	10,63

A 3. táblázat a szünetek számának és időtartamának alakulását szemlélteti a tempó változtatásának függvényében.

3. táblázat

Az olvasás tempója	Szünetek száma		Szünetek időtartama (s)	
	gyakorlott	átlagos	gyakorlott	átlagos
	beszélők			
lassú	33,4	34,9	21,08	17,10
természetes	23,2	28,0	14,93	13,10
gyors	11,8	13,1	5,49	5,71

A szünetek száma a gyakorlott és az átlagos beszélőknél nagyjából megegyezett a tempó lassításakor és a gyorsításkor, a természetes tempóban viszont az átlagos beszélőknél volt valamivel több. A gyakorlott beszélőknél a természetes tempóhoz képest a lassításkor ugyanolyan mértékben nőtt a szünetek aránya, (23-ról 33-ra), mint amilyen mértékben csökkent (23-ról kb. 12-re) a gyorsításkor. Ehhez képest az átlagos beszélőknél a lassítás kisebb mértékben növelte a szünetek számát (28-ról 35-re), mint a természetes tempóban. A gyorsításkor pedig több mint a felére csökkent a szünetek száma (28-ról 13-ra). A szünetek időtartama a két csoportban szintén hasonlóan alakult valamennyi tempóban. (A különbség a két csoport között a lassú tempóban mutatkozott legnagyobb, kb. 4



hang/s volt.) A tempó lassításakor csak a gyakorlott beszélőknél növekedett jelentősen az időtartam, míg a gyorsítás mindkét csoportban és azonos mértékben eredményezte a szünetek időtartamának csökkenését.

Megállapítható, hogy az adatközlőknél az olvasás folyamán az olvasási időnek 17,86–22,15%-át tették ki a szünetek, azaz az olvasáskor relative kevés és szinte pontosan bejósolható a szünetek aránya, mert más jellegű folyamatról van szó, mint a spontán beszédben. A szünetek időtartama a tempó változtatásakor az olvasás idejével egyenesen arányosan változott, a gyorsításkor azonban erőteljesebb volt a csökkenés. A szünetek száma viszont a gyorsításkor és a lassításkor is erőteljesen változott. Mindezek további bizonyítékai annak, hogy a tempó gyorsítása főképpen az időtartam csökkenésével, a tempó lassítása pedig a szünetek számának növekedésével jellemezhető. A gyakorlott és az átlagos beszélők olvasása nemcsak a tempót illetően, hanem a szüneteket tekintve sem mutat lényeges különbségeket.

A 4. táblázatban összefoglaltuk, hogy a szintaktikai szüneteket milyen arányban tartották az adatközlők a tempó változtatásakor.

4. táblázat

Szintaktikai határ	Szünetet tartanak az adatközlők %-ban								
	lassú tempó			természetes tempó			gyors tempó		
	I.	II.	III.	I.	II.	III.	I.	II.	III.
	csoport								
bekezdés	100	100	95	100	100	90	100	100	80
mondat	100	100	100	100	100	100	76	60	64
tagmondat	93,8	80	87,5	67,5	63,8	68,8	22,5	27,5	25
szavak	39,1	31,6	34,4	32,5	35,2	33,3	20	22,7	22,9

Amint az a táblázatból kitűnik, a szöveg tagolását jelző grammatikus szüneteket (vö. Fónagy 1967) nagyjából azonos arányban tartották a gyakorlott és az átlagos beszélők, és az sem befolyásolta ezt döntő mértékben, hogy a szöveg már tagolódott a különböző szintaktikai egységekre, vagy az adatközlőknek kellett ezt a tagolást jelölniük. A természetes tempóban a kisebb grammatikai-szintaktikai szerkezetek határán kevesebb volt a szünetek aránya. A tempó lassítása jelentős mértékben csak a tagmondatok esetében növelte a szünetek arányát, (kismértékű növekedés volt tapasztalható a szavak közötti szüneteknél is, de csak a gyakorlott beszélőknél). A gyorsítás már a mondatoknál is számottevő csökkenést eredményezett, hasonlóan a tagmondatok és a szavak közötti szünetekhez (a legerőteljesebben a tagmondatok közötti szünetek arányában mutatkozott). Az adatok egyfelől alátámasztják a szakirodalom azon megállapítását, hogy a hangos megnyilatkozások esetén a szintagmák, tagmondatok jelzésére más eszközök is vannak (Wacha 1988). Másfelől a magyar nyelvben is igazolható – hasonlóan a németben, svédben, angolban kapott eredményekhez (vö. Strangert 1991) –,

hogy a szöveg nagy szerkezeti egységeinek határán az adatközlőktől, a közpon-  
tozás meglététől vagy hiányától, valamint a tempótól független a szünetek meg-  
jelenése. Ugyanakkor a kisebb szintaktikai szerkezetek határán a szünetek meg-  
jelenésében a beszédsebesség a meghatározó. Minél lassúbb a felolvasás, annál  
gyakoribb a szünet a tagmondat-, illetve szóhatáron és fordítva.

Az 5. táblázat a szintaktikai szünetek időtartamértékeit mutatja a gyakorlott  
és az átlagos beszélők szerinti megoszlásban. (Az átlagos beszélőknél az a tény,  
hogy tagolniuk kellett a szöveget vagy nem, ezen szünetek időtartamában sem  
mutatott jelentős különbségeket.)

5. táblázat

Szintak- tikai határ	A szünet időtartama (ms)					
	lassú tempó		természetes tempó		gyors tempó	
	gyakorlott	átlagos	gyakorlott	átlagos	gyakorlott	átlagos
bekezdés	1548	761,7	1146,5	664,2	637,5	388,5
mondat	1020,5	655,8	764,2	555,4	330,8	255,4
tagmondat	599,1	405,1	261,2	266,2	159,1	138,6
szóhatár	120,8	118,7	115,7	138,6	50	97,2

A gyakorlott beszélők a szavak közti szüneteket leszámítva, csaknem vala-  
mennyi tempóban lényegesen hosszabb szünetekkel tagolták a szöveget, mint az  
átlagos beszélők. A különbségek a bekezdések közötti és a mondathatárokon tar-  
tottaknál mutatkoztak a legerőteljesebben. A természetes tempóban a következő  
összefüggés érvényesült mindkét csoportban: minél nagyobb volt a szintaktikai  
szerkezet, annál hosszabb volt az azt követő szünet. Az egyes szintaktikai szüne-  
tek között minden esetben jelentős időtartam-különbségek mutatkoztak. A tempó  
lassítása a szavak közti szünetek kivételével valamennyi szünet időtartamának  
arányos növekedését, a gyorsítás pedig mindegyiknek arányos csökkenését ered-  
ményezte mindkét csoportban. A természetes tempóhoz képest a tagmondatok  
közti szüneteket kivéve, a gyorsításkor mindig nagyobb mértékű volt az időtar-  
tam változása, mint lassításkor. Azaz, ha egy szó a grammatikai-szintaktikai  
szerkezetbe szorosan beletartozik, akkor kisebb a hírértéke, s a szoros integrált-  
ság miatt rövidebb az előtte levő szünet hossza is. Ez az összefüggés az olvasás-  
kor még fokozottabban érvényesül. Érvényességük azonban nem a lexémák vá-  
lasztásának, hanem valószínűsíthetően a szöveg értelmi, érzelmi viszonyainak  
megjelenítésére való törekvésnek a függvénye. Ez a kiemelés a mondatot alkotó  
és szorosan összetartozó tagmondatok, illetőleg szó szerkezetek esetében nem a  
szünetekkel és időtartamukkal, hanem más mondatfonetikai eszközökkel valósul  
meg.

A kötőszóknak a szünetre gyakorolt hatása a következőképpen volt jellemez-  
hető. A kötőszóval kapcsolódó tagmondatoknál a gyakorlott és az átlagos beszé-



lők is a természetes és a gyors tempóban kevesebb, a lassúban pedig több szünetet tartottak, mint az anélkül kapcsolódóknál. A szünetek minden esetben csak a kötőszavak előtt fordultak elő, szemben a spontán beszéddel (ez utóbbira ld. Szende 1979). (A százalékos adatok a kötőszóval kapcsolódó tagmondatoknál: 92,5% 63,3%, 20% a lassú, természetes és gyors tempóban, a kötőszó nélkülieknél pedig: 86,5%, 69,5%, 27%.) A kötőszó előtti szünet megjelenése jelentősen függ a kötőszó típusától is (ld. 6. táblázat).

6. táblázat

A kötőszó típusa	A szünet előfordulása %-ban					
	gyakorlott beszélők			átlagos beszélők		
	lassú	természetes	gyors	lassú	természetes	gyors
hogya	100	73,3	13,3	80	56,7	10
mert	100	60	0	70	40	10
de	100	20	0	100	60	20
ha	100	100	80	100	95	60

A táblázatból kitűnik, hogy már a természetes tempóban sem egyenlő mértékben jelennek meg a szünetek, s a feltételezésekkel ellentétben nem mindig van szünet ott, ahol például azt várnánk. (Hasonlítsuk össze pl. a *hogya* és a *ha* kötőszók előtti szünetek arányát!) A gyakorlott és az átlagos beszélők között a kötőszavak előtti szünetek arányában a kötőszó típusától függően jelentős eltérések mutatkoztak. A tempó gyorsításakor mindkét csoportban különböző mértékben tartottak szüneteket a kötőszók előtt, míg a lassítás csaknem minden kötőszó előtt szünetet eredményezett.

Az adott tagmondat grammatikai-szintaktikai szerkezete a szünetek megjelenésében a következőképpen érvényesült. A gyakorlott beszélők – a gyors tempót kivéve – az alárendelő tagmondat előtt, az átlagos beszélők pedig – ismét a gyors tempót leszámítva – a mellérendelő mondatban tartottak gyakran szüneteket. (A százalékos értékek: alárendelő mondat előtti szünetek aránya a gyakorlott beszélőknél 98%, 80%, és 24% a lassú, a természetes és a gyors ejtésben, a mellérendelő mondatban pedig: 86,6%, 46,6% és 53%. Ugyanezen értékek az átlagos beszélőknél: 75,7%, 64%, 21%, illetve 85%, 71,7% és 35%.)

Fónagy (1967) szerint, ha az alárendelő mondat megelőzi a főmondatot, akkor gyakoribb a szünetek előfordulása. Anyagunkban ez a természetes tempóban szintén kimutatható volt, a tempó változtatásakor azonban nem. Olvasáskor erősen jelentkezik az a tendencia, hogy ha hiányzik a két tagmondat kapcsolatának jellegére utaló kötőszó, akkor szünet választja el a két tagmondatot, vagyis az írásjelnek vizuális hatása a szünetre nézve erősebben érvényesül, mint a kötőszók vizuális hatása. Másfelől a kötőszó előtti szünetek megjelenése nem elsősorban a

kötőszók típusától függ, hanem attól, hogy az adott kötőszó alárendelő vagy melérendelő szerkesztésű mondatokat kapcsol-e össze.

Elemeztük a gondolatot záró pont, a tagmondatokat, szószerkezeteket elválasztó vessző, a közbevetés elhatárolását jelző gondolatjelnek és végül a kettőspontnak a szünetként való megjelenítését (7. táblázat).

7. táblázat

Írásjel	A szünet megjelenése %-ban					
	gyakorlott beszélők			átlagos beszélők		
	lassú	természetes	gyors	lassú	természetes	gyors
vessző	91,7	57,6	12,6	77,8	67,1	23,5
pont	100	100	86,6	98,9	97,8	72,2
kettősp.	100	100	0	80	80	20
közbeék.	100	70	70	95	80	35

Természetes tempó esetén a gyakorlott és az átlagos beszélők is a gondolatot záró pont után nagyjából azonos arányban, közel 100%-ban tartottak szüneteket. A többi írásjel utáni szünetek aránya a két csoportban az írásjel típusától függően különböző mértékű volt. A tempó lassítása valamennyi írásjel után gyakoribb szünetet eredményezett mindkét csoportban. A tempó gyorsításakor különböző mértékben csökkent az írásjelek utáni szünetek aránya. A gyakorlott beszélők jórészt csak a pont után, az átlagos beszélők ehhez képest a vesszőknél is és a gondolatjel után is gyakrabban tartottak szüneteket.

### Megbeszélés

Vizsgálatunkban a hangos olvasás közbeni szüneteket elemeztük, részben a tempó változtatásának függvényében, részben az adatközlőknek aszerinti megoszlása alapján, hogy ún. gyakorlott vagy átlagos beszélők. Anyagunk nem teszi lehetővé, hogy messzemenő következtetéseket vonjunk le, néhány megállapítást azonban tehetünk. Ezek a következők:

1. Az olvasás tempója elsősorban nem az 1 s alatt ejtett beszédhangok számával, hanem a szintaktikai határokon megjelenő szünetek számával és időtartamával jellemezhető. Az interpretatív közlésben a szöveg értelmi tagolása a bekezdések és mondatok határán tartott gyakori és hosszú szünetek tartásával történik. A tagmondatok, szószerkezetek és szavak esetében a tagolást a felolvasó főképpen a különböző mondatfonetikai eszközökkel valószínűsíti meg, ezért kevesebb és rövidebb ezeken a szintaktikai határokon a szünetek előfordulása.

2. Az olvasás tempójának megváltoztatása a szünetek számának és időtartamának megváltozását nem egyenlő mértékben érinti. A **lassítás a szünetek szá-**



**mának növekedésével, a gyorsítás pedig időtartamuk csökkenésével jellemezhető.**

3. Az olvasás tempója és annak változtatása független a központosítás meglététől vagy hiányától, valamint az adatközlők személyétől.

4. Olvasáskor a szöveg értelmi-érzelmi-logikai viszonyainak tükröztetése az elsődleges cél, amely jelentős mértékben megnyilvánul a szövegnek a szünetekkel történő tagolásában. Az írásjelek vizuális hatása erőteljesebben érvényesül, azonban nem bontja meg ezt az értelmi tagolást. A központosítás és a kötőszók vizuális hatása a szünetre nézve jelentős mértékben, de különböző módon jellemezhető: az írásjelek után sokkal gyakoribb a szünetek előfordulása, mint a kötőszavak előtt. A központosításnak a szünetre kifejtett vizuális hatása a tempó változtatásakor főképpen az olvasás sebességének csökkenésekor jelentkezik. Ez az értelmi tagolás alapján elvárható szünetek számát minden írásjel esetében növeli.

### Irodalom

- Boomer, D.: Hesitation and grammatical encoding. *Language and Speech* 8. 1965, 148–158.
- Butterworth, B.: A beszédszünetek adaléka. In: *A beszédmegértés és a beszédprodukció pszichológiája*. Szerk.: Pléh Csaba. Budapest 1989, 251–274.
- Fónagy Iván: Áthajlás, szünet, szerkezet. *Nyelvtudományi Közlemények* 69. 1967, 313–343.
- Fónagy Iván – Magdics Kára: *A magyar beszéd dallama*. Budapest 1960.
- Goldman-Eisler, F.: *Psycholinguistics*. London, New York 1968.
- Gósy Mária: *Beszédészlelés*. Budapest 1989.
- Gósy, M.: The perception of tempo. In: *Temporal Factors in Speech*. Ed. Gósy, M. Budapest 1991, 63–106.
- Hegedűs Lajos: On the problem of the pauses of speech. *ALINGuH* 3. 1953, 1–36.
- Kassai Ilona: Gyorsult-e a magyar beszéd tempója? (ebben a kötetben)
- Maclay, H. – Osgood, Ch. E.: Hesitation phenomena in spontaneous speech. *Word* 1. 1959, 19–44.
- Martin, J. G.: On judging pauses in spontaneous speech. *Journal of Verbal Learning and Verbal Behaviour* 9. 1970, 75–78.
- Sallai János – Szende Tamás: A szünet mint funkció. *MNyTK*. 143. 1975, 1–36.
- Strangert, E.: Pausing in texts read aloud. *Proceedings of the XIIth International Congress of Phonetic Sciences*. Vol 4. Aix-en-Provence 1991, 238–242.
- Szende Tamás: A szünet és a junktúra. *Magyar Fonetikai Füzetek* 4. 1979, 7–32.
- Váradai Tamás: A beszédsszünet szubjektív és objektív regisztrálásának összevetéséről. In: *Beszélt nyelvi tanulmányok*. Szerk.: Kontra Miklós. Budapest 1988, 44–59.
- Wacha Imre: Élő nyelvi (spontán) szövegek megnyilatkozásainak (szintaktikai) vizsgálati szempontjaihoz (a gázdagréti kábeltelevízió élő nyelvi felvételei alapján). In: *Beszélt nyelvi tanulmányok*. Szerk.: Kontra Miklós. Budapest 1988, 102–158.

**A tanulmány alapját képező kutatás az OTKA 3211 támogatásával történt.**

# **A FONETIKAI ÉS PSZICHOLINGVISZTIKAI KUTATÁSOK ÚJABB EREDMÉNYEINEK ALKALMAZÁSA A TANÍTÓKÉPZÉSBEN**

**A. Jászó Anna**

**Budapesti Tanítóképző Főiskola**

Mindnyájunk előtt ismeretes, hogy az utóbbi két évtizedben sok új eredmény született a kutatásban. Az oktatásra az jellemző, hogy egyrészt bizonyos mértékig igazodik az új tudományos eredményekhez – hol válogatva, hol új rendszerek kiépítésének az igényével –, másrészt pedig teljesen furcsa dolgokat művel, ellentmondva mind az új tudományos eredményeknek, mind a józan észnek. Mindenesetre nagyon megsínyli azt a tényt, hogy nincsenek megírva a közérthető szintézisek, nincsenek lefordítva fontos alpművek. Olykor maguk a tan-könyvszerzők is meglepődnek, mikor tudomásukra hozzuk "alapjaikat".

A tudományos kutatás és az oktatás két pólusa között elhelyezkedő tanítóképzés – és az egész főiskolai oktatás – nincsen a legjobb helyzetben: egyrészt olyan tananyagot kell kialakítani, mely korszerű és használható, másrészt fel kell készítenie a hallgatókat a tanítási gyakorlat alternatíváira. Nagyon színes ugyanis manapság a paletta: egyes olvasástanítási módszerekben szóképeket olvastatnak, a másikban hangokat/betűket tanítanak; az egyik nyelvtanítási program ún. elnyújtott ismeretszerzéssel kerülgeti a nyelvtant, és ódzkodik a terminusok korai használatától; a másik predikatív és jelzős szerkezetek egymásba történő átalakítását végezteti pici hétéves másodikosokkal; a harmadik a transzformációs generatív grammatikára és a pragmatikára építi rendszerét; s van olyan is, amelyik az ún. elemi grammatikát tanítja, annyit adagolva belőle, amennyit a gyakorlat – azaz a helyesejtés és a helyesírás – megkíván.

Tetézi a nehézséget a magyarórák számáért vívott állandó küzdelem is. 1950 óta kimutathatóan minden tantervi változtatás a magyarórák számának a csökkentésével járt. Különösen érzékeny veszteséget jelentett az, hogy 1978-ban a negyedikes órák 25%-át levették, és az oroszoknak adták. Ezt az óraszámcsökkenést azóta sem korrigálták, pedig a negyedik osztály igencsak kritikus a felső tagozatra való felkészítés szempontjából.

A következőket tehetjük ebben a helyzetben. 1. Úgy gondoltuk, hogy mindekelőtt korszerű, színvonalas, használható és megemészthető nyelvészeti anyagot kell hallgatóink kezébe adnunk. 2. Majd – ezzel a nyelvészeti anyaggal összhangban – meg kell írunk az anyanyelvi tárgyak tanításának módszertanát, mégpedig oly módon, hogy kettős feladatot teljesítsen: adjon megbízható tudást, mellyel minden körülmények között eredményesen lehessen tanítani, s ezenkívül



tájékoztasson a különféle oktatási programok anyagáról és módszeréről. A nyelvészeti anyagot publikáltuk, lényegében már hét éve használjuk. Most a módszertan kiadását tervezzük. A megírandó módszertan kapcsán csak a beszédművelésről és az olvasás-írás tanításáról kívánok szót ejteni, mivel ezek állnak közel a fonetikához; az olvasástanítás kapcsán pedig részletesen szeretném bemutatni az általam elfogadhatónak tartott pszicholingvisztikai (olvasáspszichológiai) alapot.

## 1. A nyelvészeti anyag

A tanítóképzőkben hármas célt tűztünk ki magunk elé: megtaníttuk az elméleti anyagot, begyakoroljuk, s fejlesztjük a helyesírási készséget. Ezen célok elérésére az alábbi tankönyveggyűttest írtuk meg:

*A magyar nyelv*, Tankönyvkiadó 1985, szerkesztette Adamikné Jászó Anna, írták: Adamikné Jászó Anna, Albert Sándorné, Bokor József, Hangay Zoltán, Kálmánné Bors Irén, Király Lajos, Cs. Nagy Lajos. Később jelentősen átdolgozva könyv formában is megjelent: *A magyar nyelv könyve*, Trezor Kiadó 1991.

*Magyar nyelvészeti szöveggyűjtemény*, Tankönyvkiadó 1989, szerkesztette Hangay Zoltán, összeállította: Adamik Tamásné, Hangay Zoltán, Cs. Nagy Lajos, Csikvári Gábor, Kálmánné Bors Irén.

*Magyar nyelvi gyakorlókönyv*, Tankönyvkiadó 1989, szerkesztette Hangay Zoltán, írták: Fülöp László, Kernya Róza, Szikoráné Kovács Eszter, Tomory Júlia; később könyv formájában is megjelent: *Magyar nyelvi gyakorlókönyv*, Trezor Kiadó 1992.

*Helyesírási gyakorlókönyv*, Magyar Eszperantó Szövetség 1989, írta Cs. Nagy Lajos.

Tudomásom szerint *A magyar nyelv* és *A magyar nyelv könyve* az első olyan összefoglalás, mely szövegtannal zárja le a nyelvreírás hagyományos fejezeteit; s egy nyelvtankönyv megszokott fejezetein kívül tájékoztat a névtanról, a nyelvjárásokról – a második változat a határokon túli nyelvjárásokról is –, a gyermeknyelv elsajátításáról és jellemzőiről, valamint a modern nyelvészeti irányzatokról. Most nincsen mód a program részletes ismertetésére, ezt különben is megtette Hangay Zoltán: *Nyelvészeti munka a Budapesti Tanítóképző Főiskolán c. tanulmányában* (Hangay 1988; ld. még Szikszai 1988; 1989). Nehéz volna felsorolni azokat az újításokat, melyeket a szerzők az újabb szakirodalom alapján beépítettek az egyes fejezetekbe: tehát a könyvnek nemcsak a globális szerkezete új, hanem sok részlete is. Bár a nagyközönség is haszonnal forgathatja, elsődleges célja a tanítójelöltek felkészítése. A mai iskolai gyakorlatban ugyanis többféle programot használnak. Ezt a tényt megfontolva tartalmazza a könyv a külön-

féle nyelvészeti irányzatok ismertetését, alapot kíván nyújtani a különféle szemléletű nyelvtanok tanításához.

A következőkben csak a hangtani fejezetet ismertetem. Megírásakor elsősorban Vértés O. András és Gósy Mária munkásságára támaszkodtam. Vértés O. András munkássága a fonetikának csaknem minden területéhez kapcsolódik, tőle vettem át például az evokatív és az emfatikus variáns fogalmát; a beszédhibák és a kiejtési hibák leírása, a beszédhangok expresszivitásáról szóló fejezet az ő dolgozatain alapul. Gósy Mária tanulmányai a beszédpercepcióról szóló fejezet alapját képezik. Természetesen sok egyéb eredményt is beépítettem a hangtani fejezetbe, sok-sok részlet az új kutatásokról is tájékoztatja az olvasót.

A hangtani fejezet igazi újdonságát – ha szabad azt mondani, hogy eredetiségét – nem részleteiben, hanem inkább megkomponálásában látom. Figyelembe vettem ugyanis Bolla Kálmánnak irányadó tanulmányát (Bolla 1982). Ebben Bolla a beszédlánc szakaszainak megfelelően a hangtan alábbi területeit állapítja meg: kortikális, artikulációs, akusztikai, auditorikus és percepciós fonetika. Ezen felosztást követve az eddigieknél kissé nagyobb teret kapott az akusztika, s az újabban írt tananyagban először szerepel a beszédprodukcióval és a beszédpercepcióval foglalkozó fejezet. Ezzel az újítással a hangtan összhangba kerül a kommunikációelméleten nyugvó bevezető fejezettel; továbbá a percepciós fonetika alapul szolgál a beszédművelés-beszédfejlesztés és az olvasástanítás módszertanához. A beszédhibákról, a kiejtési hibákról és a beszédhangok expresszivitásáról szóló fejezetek is a módszertant kívánják megalapozni.

## 2. A módszertani anyag

Az a célunk, hogy a megírandó módszertan maximális mértékben támaszkodjék *A magyar nyelv könyvében* megfogalmazott elméleti anyagra. Jelenleg az a helyzet a tanítóképzésben, hogy az egyes részterületek szétesnek. A módszertan erősen általános pedagógiai szempontból van megírva – mint azt a címében szereplő tantárgypedagógia elnevezés is mutatja –, kevés, a gyakorlatban használható konkrétumot tartalmaz. Nem is igen használják, sem a főiskolai módszertanok, sem a tanítási gyakorlatokat vezető tanítónők. Valamiféle a tapasztalaton nyugvó konszenzus viszi előre és tartja működésben az egész apparátust, s ez a gyakorlat messze esik mindattól, ami az utóbbi évtizedekben a kapcsolatos szakmában lezajlott.

Pillanatnyi utasítások végrehajtásán nyugszik a gyakorlat, akár hagyományos, akár ún. modern. Éppen ezért tartjuk fontosnak egy új módszertan megírását, egy olyan módszertanét, mely saját szakmai alapjain nyugszik, azaz a nyelvészetén. (Ez nem jelenti azt, hogy nem vesszük figyelembe a segédtudományok: a filozófia, a pedagógia és a pszichológia szempontjait.)



A megírandó módszertan a következő fejezetekből fog állni (egyelőre különálló tankönyveket kívánunk írni, melyeket később valószínűleg egybeszerkesztünk): A kommunikációs képességek fejlesztésének a tanítása (beszédművelés, beszédfejlesztés és a nem nyelvi kommunikációs képességek fejlesztése); Az olvasás-írás tanítása; A nyelvtan-helyesírás tanítása; A szövegelemzés tanítása; A fogalmazás tanítása; A dramatikus játékok tanítása. Három kiegészítő anyagot tervezünk még ehhez a törzsanyaghoz: Beszédművelési feladatgyűjtemény; Anyanyelvi módszertani szöveggyűjtemény (ebben szeretnénk egy-két jelentős külföldi tanulmány fordítását is közölni); Anyanyelvi módszertani munkafüzet tanítóképzős hallgatók számára.

A továbbiakban a beszédművelést és az olvasás-írást ismertetem.

*A beszédművelés módszertana.* Itt két egységet kívánunk elkülöníteni, mégpedig a beszéd és a beszédértés tanítását, az angolszász szakirodalomban szokásos *speaking* és *listening* mintájára. Az eddigi tantervekben és módszertanokban is volt szó a beszédértés fejlesztéséről, fontosságát mindig megemlítették, de kidolgozva még sosem volt.

A beszédművelés tervezett felépítése a következő:

- A beszédprodukció és a beszédértés fejlesztésének fontossága;
- A gyermek nyelvének fejlődése (rövid áttekintés);
- Az iskolába lépő gyermek kiejtésének a jellemzése;
- A gyermek beszédészlelésének a fejlődése;
- Az iskolába lépő gyermek beszédészlelésének a jellemzése;
- A rejtett tanterv: a tanító beszédviselkedése és hatása;
- A kiejtéstanítás tantervi céljai, feladatai és követelményrendszere;
- A kiejtéstanítás és az olvasástanítás-nyelvtanítás kapcsolata;
- A kiejtési gyakorlatok típusai: drillek és szöveghez kapcsolt gyakorlatok;
- A kiejtési gyakorlatok fokozatos felépítése;
- A kiejtési gyakorlatok beépítése a tanítási órába (órarészletek és javasolt tevékenységek);
- A beszédhibák felismerése, a beszédhibás gyerekekkel való bánásmód;
- A kiejtési hibák felismerése és javítása;
- A memoriterek fontossága a kiejtés tanításában;
- A nyelvjárási kiejtés és a hozzá való viszonyulás;
- A beszédértés fejlesztésének tantervi céljai, feladatai és követelményrendszere;
- Beszédfeldolgozási stratégiák;
- A beszédfeldolgozási gyakorlatok kapcsolata az olvasástanítással, a drámatanítással és a gyermekirodalommal;
- A beszédfeldolgozási gyakorlatok típusai;
- A beszédfeldolgozási gyakorlatok fokozatos felépítése;

A beszédfeldolgozási gyakorlatok beépítése a tanítási órába (órarészletek és javasolt tevékenységek);

A beszédészlelés vizsgálata, fejlesztési javaslatok.

Mindegyik fejezetbe beleértendő mind a szegmentális mind a szupraszegmentális tényezők fejlesztése.

Látható, hogy *A magyar nyelv könyvének* hangtani fejezete – beszédprodukcio-beszédpercepció – és a módszertan beszédművelési fejezetének a terve egymásra épül. Nagy segítséget jelent Gósy Mária *Beszédészlelés* c. könyve, (1989a); hasonlóképpen jól tudjuk használni Óvónői *Beszédviselkedés* c. kézikönyvét (Gósy 1990); valamint az általa kifejlesztett beszédészlelést és beszédmegértést vizsgáló GMP-tesztet (1989b). Ez utóbbinak működését gyakorlóiskolánkban is próbáltuk.

A beszédészlelésnél sokkal kidolgozottabb a kiejtéstanítás módszertana, elegendő Hernádi Sándor, Montágh Imre, Zsolnai József és még sokak munkásságára utalni. A legújabb szakirodalomból Fekete László *Magyar kiejtési szótárát* (1992), említtem meg, ugyancsak ő fogja összeállítani a *Beszédművelési feladatgyűjteményt*, mely egyik alapját fogja képezni a beszédművelés tanításával foglalkozó fejezetnek.

*Az olvasás-írás tanításának módszertana.* Bármilyen furcsán hangzik, nincsen jelenleg a tanításág kezébe adható, haszonnal forgatható olvasás- és írástanítási módszertan. Természetesen, vannak kézikönyvek a forgalomban lévő programokhoz, létezik a *Melyiket válasszam?* c. tanácsadó kötet (1990), de mindezek nem pótolják a szemléletet formáló és a gyakorlatot szolgáló módszertant. Van Anyanyelvi tantárgypedagógia jegyzet (Nagy J. 1983), de ebben 354 oldalból kb. 50 oldal foglalkozik az olvasás-írás tanításával, s ebből csak 10 (!) jut a kezdeti olvasástanítási módszerekre.

Magában a tanítási gyakorlatban két nagy probléma észlelhető (a számos kisebbről most nem beszélve). Az 1978-as tanterv kimondta, hogy az írás eszközjellegű tevékenység. Ez általában igaz, de a tanítás kezdő szakaszában nem így van: az írás ekkor még cél, hiszen a kisgyerek még nem tud írni, most tanulja az írást. Ezzel a kijelentéssel, valamint az idő előtti tempógyorsítás erőltetésével, a szabálytalan betűalakítások korai megengedésével elérték azt, hogy az új nemzedék írásának minősége hihetetlenül leromlott, olyannyira, hogy már-már akadály a kommunikációnak; s az írástanítás módszertana elfelejtődött. Újra kell teremteni.

A másik gond az olvasástanítási módszerek közötti eligazodás kérdése, a magyar nyelv szerkezetéhez és a magyar helyesírás sajátosságaihoz illeszkedő optimális olvasástanítási módszer kiválasztása. Külön feladat az olvasástanítás pszichológiai megalapozása, a hazai és a külföldi pszicholingvisztikai és olvasáspszichológiai kutatások eredményeinek a figyelembevételével. Az olvasás-



pszichológiai megalapozásnak sokan nagy jelentőséget tulajdonítanak. Magam is elismerem fontosságát, bár egy kissé skeptikus vagyok: annyi az olvasáspszichológiában az egymásnak ellentmondó eredmény, hogy pillanatnyilag megbízhatóbbnak tartom az olvasástanítás történetének a tanulságait. A módszertan végül is nagyrészt a tapasztalatok összegezése, s az olvasástanítás története a tapasztalatok tárháza: kiolvashatjuk belőle, hogy milyen módszerek működtek jól a múltban, s milyenek hullottak ki az idő rostáján. Egy módszertani tankönyvnek alapvetően módszertani jellegűnek kell lennie, a pszichológia és a filozófia itt "csak" a segédtudományok szerepkörét vállalhatja el, semmiképpen sem nyomakodhatnak az előtérbe (ezzel a megállapítással csak a hazai módszertan nagy mesterét, Kármán Mórt követem, egyáltalán nem mondok új dolgokat). Az olvasáspszichológiai álláspont kifejtése bonyolult feladat, ezért a dolgozat második részében külön kívánok vele foglalkozni.

A készülő módszertanban – melynek írására magam vállalkoztam – új lesz az olvasáspszichológiai és -történeti alapozás, az előkészítés részkészségeinek aprólékos leírása és tanítása (ez a fejezet összefügg az olvasási nehézségekkel és javításukkal foglalkozó fejezettel), egy optimális modell megállapítása, a jelenleg forgalomban lévő programok ismertetése, sok óravázlat és megoldási javaslat, a szövegértő olvasás fejlesztésének a kimunkálása, a tantárgyi és a felső tagozatos olvasástanítás, az olvasási és írási gondokkal küszködő gyerekekkel való foglalkozások módszere.

Az olvasás-írás tanításának a tervezett felépítése a következő:

#### I. Az olvasásról általában

Az olvasás fontossága; Az olvasás fogalma; Az olvasás fajtái;

Az olvasás motivációi; Az olvasástanítás területei;

#### Az olvasási folyamat

Az olvasási folyamatot ábrázoló modellek; Az olvasási folyamat kialakulása; Az olvasás mint készség; A betűk, a szavak és a mondatok felismerése; A gondolkodás fejlettsége és az olvasás; Metakogníció és nyelvi tudatosság; A sémaelmélet

#### Az olvasástanítási módszerek

Az olvasástanítási módszerek felsorolása; Az olvasástanítási módszer és az anyanyelv típusa; Az olvasástanítási módszer és a helyesírási rendszer; Az olvasástanítási módszerek és a pszichológiai kutatások; Az olvasástanítási módszerek és a filozófiai iskolák;

#### Az olvasástanítás és az olvasáskutatás története

#### Az optimális olvasástanítási modell

#### Az olvasás tanításának az előkészítése

Az alapozó részkészségek fejlesztése

A közvetlen részkészségek fejlesztése

- A nyelvi tudatosság fejlesztése; A szótagolás; Egy sajátos előkészítés: a pálcikás módszer; Az előkészítés a különféle programokban; A tanító munkája az előkészítő szakaszban; Óravázlatok az előkészítéshez;
- A hangtanítás és a betűtanítás menete
  - Óravázlatok a hang- és betűtanításhoz;
  - Az összeolvasás
  - Az olvasástechnika fejlesztése
  - Az olvasási hibák és javításuk
  - A szövegértő olvasás fejlesztése
  - Az olvasókönyvek
  - Az olvasás a különféle tantárgyakban
  - Az olvasás tanítása a felsőbb osztályokban
- II. Az írás tanítása
  - Az írásról általában
    - Az írás fontossága; Az írás fogalma; Az írás fajtái; Az írás motivációi; Az írástanítás területei;
  - Az írás folyamata
  - Az írástanítási módszerek
  - Az írástanítás története
  - Az írástanítás párhuzamosságának, illetőleg késleltetésének a kérdése
  - Az optimális írástanítási modell
  - Az írás tanításának az előkészítése
    - Az alapozó részkészségek fejlesztése
    - A közvetlen részkészségek fejlesztése: betűelemek írása és kapcsolása;
  - A betűk alakítása
    - Óravázlatok a betűk írásához
  - A betűk kapcsolása
  - Az íráskészség fejlesztése
- III. A jelenleg használt olvasás- és írástanítási programok
  - A hangoztató-elemző-összetevő módszer, globális előprogrammal
  - A globális módszer
  - Az intenzív-kombinált módszer
  - Az olvasástanítás a nyelvi-irodalmi-kommunikációs programban
  - A heurisztikus (programozásos) módszer
- IV. Az olvasási-írási gondokkal küszködő gyerekek tanítása
  - Az olvasási nehézségek és a diszlexia; a diszgráfia
  - Tesztek az olvasási nehézségek megelőzésére: GMP, Inizán, DPT
  - Megelőző munka az óvodában
  - Program diszlexiás gyerekeknek
  - Program a kisegítő iskolák számára
  - Felső tagozatos korrekciós programok



Ezzel a könyvvel szeretnénk egy megbízható módszertant a tanítóság kezébe adni, melyet követve biztonságosan meg lehessen tanítani az írást és az olvasást. Azt is szeretnénk elérni, hogy tisztán lássák: az olvasástanítás az egész iskolai munka megalapozása. Ez nemcsak azt jelenti, hogy az a gyerek tud tovább haladni tanulmányaiban, aki jól tud olvasni; hanem azt is, hogy az elsős írás- és olvasástanítás magában foglalja a többi anyanyelvi részterületet, így például az olvasás-írás tanításának a minőségétől függ a helyesírás későbbi minősége is. Ezért végtelenül fontos az első iskolaév és az elsős tanító tudása, személyisége.

### 3. A pszicholingvisztikai (olvasáspszichológiai) alapok

A nyelvi tudatosság (language awareness) megléte vagy hiánya kulcskérdés az olvasástanulásban, különösen az olvasástanulás előkészítő szakaszában. Azt a kisgyereket lehet könnyen megtanítani olvasni, akinek van némi fogalma a nyelv szerkezetéről, a mondatokról, a szavakról, a szótagokról és a hangokról. Következésképpen az előkészítésnek igen nagy jelentősége van az olvasás és – hangsúlyozzuk – a helyesírás készségszintű elsajátítása szempontjából. A helyesírás idevonása nem mindenki számára természetes. Vannak ugyanis olyan olvasáskutatók, akik azt vallják, hogy bármelyik módszerrel előbb meg lehet tanítani a gyereket olvasni. Ez talán igaz, de a helyesírási készség kialakulásának nem kedvez mindegyik módszer, csak az, amelyik törődik a nyelvi tudatosság fejlesztésével, vagyis tartalmaz előkészítő szakaszt. Azt felejtik el sokan, hogy az olvasástanulás során mindegyik anyanyelvi részterület alapozódik, nemcsak az olvasási készség.

A továbbiakban ismertetem 1. az újabb olvasáspszichológiai irányzatokat és a rájuk épülő gyakorlati megoldásokat, 2. a nyelvi tudatossággal kapcsolatos újabb kísérleteket, 3. a tapasztalatból – a tantárgytörténetből – levonható tanulságokat, 4. végezetül leírok egy optimális olvasástanítási modellt, mégpedig a legújabb olvasáspszichológiai kísérletekre és a felhalmozott tapasztalatokra alapozva.

**1. Az újabb olvasáspszichológiai irányzatok.** Az utóbbi évtizedek – az 1970-es és az 1980-as évek – olvasáspszichológiai kutatásaiban két irányzat figyelhető meg. Az egyik nem tesz különbséget a nyelv elsajátításának és az olvasás elsajátításának folyamata között; a másik azt tanítja, hogy a nyelv elsajátítása más folyamat, mint az olvasás elsajátítása.

Az első irányzathoz olyan amerikai pszicholingvisták tartoznak mint *Kenneth Goodman* és *Frank Smith*. Ők azt tanítják, hogy a nyelv elsajátítása és az olvasás elsajátítása (nem is beszélnek olvasástanulásról) azonos folyamat. A kisgyerek úgy tanul meg olvasni, ahogyan beszélni. A beszéd létrehozása, a beszédprodukció aktív folyamat. Beszéléskor a felszínre hozzuk tudatunk mélyéről a mondatokat a velünk született nyelvi kompetenciánk segítségével. A beszédértés is aktív folyamat, ez is nyelvi kompetenciánk segítségével, elsősorban szemantikai



asszociációk működtetésével történik. Hasonló folyamatok zajlanak le tudatunkban az olvasott szöveg megértésekor. Goodman szerint a szöveg megértése nem a felszíni struktúrában létezik, ezért pusztá dekódolás segítségével nem ragadhatjuk meg, hanem a tapasztalatok során felhalmozódott nyelvi tudásunkkal hozzuk a felszínre. Az olvasónak az író szándékát kell kitalálnia; szakadatlan interakció van a két fél között, mint egy teniszjátzmában. Ezért mondja Goodman, hogy az olvasás: *pszicholingvisztikai találósjáték* (Goodman 1970). Mivel nem a dekódolás a lényeg, az olvasást nem kell külön tanítani; csak jó szöveget kell a gyerekek elé tenni, és segítenünk kell őket a jelentés kihámozásában. A Goodman-féle elmélet alapján álló módszertanok egyik kulcsszava a **segítség** (help). Látható, hogy Goodman tanai mögött *Noam Chomsky* nyelvelmélete áll, de megítélésem szerint sajátos interpretációban: Goodman ugyanis azonosítja a jelentést a mélyszerkezettel (ez a nézet eléggé elterjedt az amerikai olvasási szakirodalomban<sup>1</sup>).

A Goodman-féle irányzatot többen bírálták, például *John Downing* mondván, hogy a Goodman által megadott modell a felnőt, a gyakorlott olvasóra jellemző, nem pedig a kezdőre. Meglátásom szerint is így van.

Minden kritika ellenére a Goodman-féle pszicholingvista álláspont igen népszerű volt a hetvenes években Angliában is, olyannyira, hogy hívei teljesen elvetették a hang- és betűtanítást, sőt: igen maradónak tartották. Újabban az USA-ban, Ausztráliában, Új Zélandon<sup>2</sup> új néven jelentkezik a hetvenes évek irányzata: *egész nyelvi megközelítésnek* hívják (whole language approach). Nem betűket, nem szavakat, nem is mondatokat tesznek a kezdő gyerek elé, hanem az egész nyelvet, azaz szövegeket. A gyermek dolga a felfedezés: *felfedező olvasásról* (invented reading) és *felfedező írásról* (invented writing) beszélnek. Például egy kisgyermek átmenetileg írhatja a *pencil* helyett azt, hogy *pesl*, a *water* helyett *wotr*, a *think* helyett *theek*, a tanár ennek örül, mert a gyerek íráskísérletében a fonológiai tudatosság ébredését üdvözli. Majd az évek során kialakul a helyes forma, nem kell sürgetni.

A tanár feladata a gyermek figyelése – a *kid-watch* is kulcsszó –, feljegyzések készítése, és természetesen a nagy felfedezés vezetése-irányítása. Az újabb publikációk tele vannak a feljegyzések vezetésére és a szervezésre vonatkozó tanácsokkal.<sup>3</sup> Olyan a tanár, mint egy pszichológiai kísérlet vezetője, ujjong, ha a hét-nyolc éves gyerek végre jól-rosszul eltalálja egy-egy szó helyesírását. Természetesen érdekes dolog nyomon követni a gyermek gondolkodásának a fejlődését, de kérdéses, lehet-e az iskolából kísérleti laboratóriumot csinálni.

Ha megismerjük egy kissé az amerikai olvasástanítás történetét, megértjük az egész nyelvi mozgalmat. Azt kell tudnunk, hogy az egész-szó-módszerre (hazai terminológiával: globális módszerre) épülő olvasókönyvek, a *basal reader*-ek tudományosan összeállított szólistákon alapuló igencsak mesterkélt mondatkákat, szövegeket közöltek és közölnek ma is. Merevek, érdektelenek, s a hozzájuk kapcsolódó tesztelés is merev. A tanár egyéniségét is megölik, nemcsak a gyerekek



érdeklődését. Érthető, hogy a bírálók az érdekes szövegekhez fordultak, és a jó gyerekirodalmat állították a középpontba. Felfedezték (!) az irodalmi olvasókönyveket. Nem alkalmaznak tesztet, s engedik a gyerekeket saját tempójukban haladni. A gyerekek folyamatos megfigyelésére való buzdítás valószínűleg rugalmasabbá, intelligensebbé teszi a tanárokat, s egyben felelősségteljesebbé is. Valójában nehéz az egész nyelvi mozgalomról ítéletet alkotni, az amerikai körülményeket figyelembe véve sok pozitívuma is van.

A *másik irányzathoz* a nyelvész Ignatius Mattingly és a pszichológus John Downing tartozik. Ők azt tanítják, hogy a nyelv elsajátítása és az olvasás megtanulása két különböző folyamat. Különbözik egymástól a beszéd megértésének és a nyomtatott szöveg megértésének a folyamata. Az olvasás nyelvi tudatosságot kíván.

A *nyelvi tudatosság* fogalmát Ignatius Mattingly vezette be egy 1972-ben írt tanulmányában (1972). Abból indult ki, hogy lényeges különbség van a beszéd és az olvasott szöveg megértése között. A beszédet és a beszédértést elsődleges nyelvi tevékenységnek nevezi. Ezzel szemben az olvasás másodlagos nyelvi tevékenység, szinte parazita, mert az elsődleges tevékenységen élősködik. Az olvasás nyelvi tudatosságot kíván. Érvelésének e pontján vezette be Mattingly a *linguistic awareness*, tulajdonképpen nyelvészeti tudatosság fogalmát.<sup>4</sup>

Őt is többen támadták (Chomsky 1972), talán félreértve a problémát. Mattingly ugyanis nem terminológiai tudatosságról beszél, nem nyelvtani, metanyelvi ismeretről, hanem csupán bizonyos nyelvi elemek felismeréséről, elkülönítéséről. Egy későbbi, 1984-es tanulmányában (Mattingly 1984) a tudatosság (consciousness) terminus helyett a *hozzáférhetőség* (access) terminust használja.

Mattingly kétféle agybeli tevékenységet különböztet meg az olvasás kapcsán. Az egyik az *analitikus olvasás*, melynek során az olvasó a mondat leírt szavait azonosítja az agyában tárolt lexikonnal, grammatikai analízist végez, azaz elemeire bontja a szót; s ezen műveletek végzésével párhuzamosan megérti a mondatot. A másik az *impresszionisztikus olvasás*, melynek során az olvasó megpróbálja kitalálni a mondat jelentését, anélkül, hogy grammatikai analízist végezne. Az effajta olvasás a szó ismerős formájára támaszkodik, és jelentésbeli asszociációk hálóját működteti. Az analitikus olvasás munkaigényesebb és lassabb, a kezdőre jellemző. A gyakorlott olvasó olvas impresszionisztikus módon, tempója a szöveg minőségétől és a motivációtól függ. Mattingly úgy véli, hogy nem igazán olvasó, aki képes a mondat analízisére, ha szükséges; azaz nem képes analitikusan olvasni, ha a szükség úgy kívánja. Mattingly két érveléssel támogatja meg gondolatmenetét: 1. az ortográfia jellegével, 2. a percepcióval foglalkozó kísérletekkel.

1. Az ortográfia megmutatja az olvasási folyamat analitikus jellegét. Mindig nyelvészeti átírás, mindig morfofonetikus<sup>5</sup> jellegű, vagyis az alfabetikus rendszerek nemcsak a fonémák, hanem a morfémák állandóságára is törekszenek. Az ol-



vasó a fonémákat és a morféákat ismeri fel. Közben az impresszionisztikus olvasás is működik a szemantikai asszociációkkal. De ha az olvasó nem ismeri a szó fonéma-morféma felépítését, bizonyos esetben nincsen más választása, mint találgatni. Ilyenkor történnek a helyettesítések. Tehát szükség van egy olyan grammatikai tudásra, melyet az olvasás során feltétlenül alkalmazni kell, de nem kell feltétlenül alkalmazni a beszédértés során.

2. A beszédpercepció kísérletekkel foglalkozó Liberman és társai azt bizonyították be, hogy a rövid távú memóriában a hangok tárolódnak, következőképpen olvasáskor mindig fonetikai dekódolás történik (Liberman 1967).

Az a gyerek, akinek nyelve normálisan fejlődik, kifejleszti az olvasáshoz szükséges grammatikai tudást. Az a gyermek, akinek nyelvi fejlődése nem megfelelő vagy megáll, nehézségekkel fog küszködni; ilyenkor kell különleges instrukciót alkalmazni, vagy pedig meg kell változtatni a gyermek környezetét.

John Downing, aki számos alapmű szerzője, továbbfejlesztette Mattingly elméletét (1973). *Feladattudat az olvasási készség fejlődésében* (Task Awareness in the Development of Reading Skill)<sup>6</sup>, valamint vaskos *Olvasáspszichológiájában* azt fejti ki, hogy az olvasás készség. A pszichológia foglalkozik a készséggel mint viselkedési formával. Az olvasás pedig beleillik a készség elnevezésű viselkedési formába. Az olvasáspszichológia társszerzőjével, C. K. Leonggal elemezték a pszichológiai készség-fogalmat, húsz ismérvét állapították meg, majd azt fejtették ki, hogy az olvasásra is jellemzők ezek az ismérvek, tehát az olvasás készség. Az alábbiakban a teljesség igénye nélkül soroljuk fel ezeket az ismérveket, kifejtésük ugyanis oldalakat venne igénybe.

Röviden azt mondhatjuk, hogy a *készség* valamilyen céllal végzett hasznos, komplex tevékenység, mely gyakorlást igényel. Talán legfontosabb ismérve a komplex jelleg, vagyis *részkészségekből* tevődik össze. Fontos a sok-sok részkészség *integrációja*, mégpedig *sima*, zökkenésmentes, olajozott egybekapcsolódása. Fontos az *időzítés*, nem annyira a gyorsaság, hanem a rugalmas időkezelés. A gyakorlott tevékenység *automatikus*, a gyakorlott ember fel tudja szabadítani figyelmét egyéb tevékenységre is: pl. a gyakorlott olvasó jobban tud ügyelni közönsége reakcióira. A gyakorlott embert a *cselekvésre kész állapot* jellemzi, s készenléti állapotra jellemző az *anticipáció* vagyis az elvárás. A gyakorlott olvasóban nyelvi szinten elvárások és következtetések működnek. A gyakorlott olvasó nem böngész végig az egész szóalakot, elegendő a számára néhány betű felfogása, vagy a címet látván képes következtetni a tartalomra. A gyakorlott olvasó *képes a tudatos kontrollra*, képes a *visszacsatolásra*, vagyis a javításra, továbbá *szét tudja választani a fontos és a nem fontos információkat*, pl. el tud vonatkoztatni a betű színétől, nagyságától.

A beszédpercepció kutatásokban *felismerési kulcsnak* (clue) nevezik azt a jelenséget, ami lényeges az információ átadása és megértése szempontjából. Például egy fonéma esetében a felismerési kulcsok a fonéma releváns jegyei. Az ol-



vasás esetében ilyen felismerési kulcs az illusztráció, a betű, a szó alakjának az egésze, s végső soron maga a gondolat lényege. *A gyakorlottság növekedésével a felismerési kulcsok egyre inkább belsők, tartalmiak lesznek.* A kezdőnek – főleg a globális módszerekben – eleinte több szüksége van az illusztrációkra a tartalom kitalálásakor. A kezdő hangosan olvas, mert szüksége van a hangos beszéd megerősítésére, majd mormol, majd már csak a szája mozog, s végül már csak a hangszálak jeleznek némi maradványt a hangos tevékenységből (ezt a jelenséget hívják szubvokalizációnak, s eszközökkel is tudják mérni). A gyakorlott olvasó fokozatosan elhagyja ezeket az ún. külső kulcsokat. Sokáig úgy gondolták, hogy gyorsabban érnek el jobb eredményt, ha kiküszöbölik a külső felismerési kulcsokat, vagyis azonnal a néma olvasást szorgalmazzák. A tapasztalat azonban azt mutatta, hogy az olvasónak át kell mennie a hangos olvasás fázisain. *A percepció egységei nagysága a gyakorlottsággal párhuzamosan növekszik.* A kezdő szótagokat, szavakat olvas, s fokozatosan lesz képes a szólamok átfogására. A gyakorlott olvasó gondolati egységeket, azaz szólamokat olvas. *A gyakorlott ember képes a stresszhelyzetben is jó teljesítményt nyújtani,* tehát nem romlik a teljesítménye akkor, ha egy osztály vagy egy idegen tanár előtt olvas.

Tehát megállapítottuk, hogy az olvasás készség. Ezért elsajátításához – mint minden egyéb készség elsajátításához – három szakasz szükséges: *egy megismerési, egy elsajátítási és egy automatizálási szakasz.* Az első szakaszban a kezdő megérti a feladatot. A kezdő olvasó esetében ez a legfontosabb szakasz, mert itt van a legtöbb részkészség. Ez az a szakasz, melyet előkészítésnek nevezünk. Ekkor kell tiszta ismereteket nyújtani, többek között tisztán kell látni a hang-betű megfeleléseket. Ez a helyesírás szempontjából is fontos.

A jó olvasó számára a hang-betű megfelelések világosak, ellenkező esetben zavar támad. Ezt a zavart M. D. Vernon megismerési zavarnak, vagyis *cognitive confusion*nak nevezte (1957). Az ő régi terminusára utalva alkotta meg Downing a *megismerés világosságának az elméletét*, azaz a *cognitive clarity theory* (Downing et al. 1983). Ezen ő a Mattingly-féle nyelvi tudatosságot érti. Csakis a tudatos tanuló akar elérni egy célt, egy bizonyos szintet. Ezt az állapotot nevezi Downing *feladattudatnak*, *task awareness*nek. Kétféle feladattudatot különböztet meg: a) *az olvasás céljának a tudatát*, b) *a nyelv szerkezetének az ismeretét*, vagyis a már sokat emlegetett nyelvi tudatosságot (linguistic awareness). Egy tesztet is kifejlesztett a kétféle tudatosság mérésére.<sup>7</sup>

2. *A nyelvi tudatossággal kapcsolatos újabb kísérletek.* 1979-ben Kanadában egy szemináriumot szerveztek a nyelvi tudatossággal kapcsolatos kutatások áttekintésére, mert ezek időközben nagyon felszaporodtak. 1984-ben John Downing és Renate Valtin, az akkor még nyugat-berlini Freie Universität tanára egy kötetet szerkesztett *Nyelvi tudatosság és olvasástanulás* címen (Downing – Valtin 1984).



A kutatók a következő kérdéseket állapították meg: 1. a nyelvi tudatosság következménye az olvasástanulásnak, fokozatosan alakul ki az olvasás-írás elsajátítása során. Renate Valtin szerint ez az álláspont olvasható ki a kötet tanulmányainak többségéből, s e véleményhez csatlakozik maga is a kötetet záró tanulmányaiban. 2. A nyelvi tudatosság előfeltétele az olvasástanulásnak. Ez Ignatius Mattingly véleménye. 3. A harmadik vélemény szerint kölcsönhatás van a nyelvi tudatosság és az olvasástanulás között, de a kezdet kezdetén feltétlenül szükséges némi nyelvi tudatosság és feladattudat. Ez John Downing véleménye. 4. Végezetül létezik olyan álláspont is, mely szerint mind a nyelvi tudatosság, mind az olvasástanulás egy harmadik, kognitív tényezőtől függ: megfelelő szellemi érettségtől és intelligenciától. Vannak olyan kutatók, akik a kölcsönhatás elméletét tartják a legvalószínűbbnek, de végül is azt is beláthatjuk, hogy ezek az álláspontok nem zárják ki egymást (ez Alfred Lie véleménye, l. később). Az kétségtelen, hogy az olvasás-írás tanulása elősegíti a nyelvi tudatosság fejlődését, feltétlenül magasabb szinten van az olvasás-írás megtanulása után, de az is tény, hogy némi előzetes ismeret és feladattudat is szükségeltetik, s ez szempontunkból igen lényeges.

E vitás kérdés – az előzmény/következmény körüli viták – eldöntésére vállalkoztak a skandináv kutatók. Két nagy anyagon végzett több évet felölelő vizsgálat eredményeit közlik az IRA (International Reading Association, Nemzetközi Olvasástársaság) tudományos folyóiratában, a *Reading Research Quarterly*-ben. Az első 1988, a második 1991 nyarán jelent meg.

*Ingvar Lundberg*nek és társainak "Egy intenzív program hatása óvodás korú gyermekek fonológiai tudatosságának ösztönzésére" c. tanulmánya (1988) válaszképpen született Renate Valtinnak az 1984-es kötetében kifejtett nézetére. A svéd és a dán szerzők azt hangsúlyozzák, hogy a hangok előzetes tudatos ismerete kritikus készség az olvasás elsajátításában. Azt bizonyították be, hogy előzetes tréninggel lehet ezt a készséget fejleszteni, továbbá hogy *ez a készség pozitívan hat mind az olvasás, mind a helyesírás elsajátítására.*

Egy nagyszabású longitudinális kísérletet folytattak le mintegy 400 gyerekkel, a 2. év végéig követve teljesítményüket. A kísérlet meglehetősen elszigetelt dániai szigetek óvodáiban és iskoláiban zajlott le. Folyamatosan tesztelték a gyerekeket (a kísérlet mérésmethodikai szempontból kifogástalan, itt nem ismertetem részletesen). A gyerekek naponta 15–20 perces tréningeket kaptak játékos formában. A gyakorlatokat fokozatosan építették fel. Az első hónapban csak általános beszédészlelési játékot adtak, beszédhangok és egyéb hangok megkülönböztetését stb. Később mondatokat bontottak szavakra. A második hónapban szótagoltattak (minden vizsgálat egyértelműen kimutatta, hogy a szótagokra bontás a legkönnyebb a gyermek számára), a harmadik hónapban kezdték csak el a hanganalizálási gyakorlatokat. Csak a kísérleti csoport kapta ezeket a tréningeket, a kontroll csoportot a Dániában szokásos érzelmi-esztétikai jellegű programmal



tanították. A két csoport közötti különbség nem is annyira az első osztályban, hanem a másodikban mutatkozott meg. A különbség a helyesírás esetében még nagyobb volt, mint az olvasás esetében: a kísérleti csoport jobb teljesítményt nyújtott mind olvasásban, mind helyesírásban.

A szerzők tehát azt hangsúlyozzák, hogy a hang tudatosítása elvégezhető az olvasástanítás megkezdése előtt és attól függetlenül továbbá azt is megállapítják, hogy a hangok előzetes ismerete megkönnyíti az olvasás elsajátítását, határozottan okozati összefüggés állapítható meg. Nagyon fontosnak tartom az olvasási és a helyesírási készség párhuzamos vizsgálatát és kapcsolatba hozását, tudniillik véleményem szerint a helyesírási készség csakis az olvasási készséggel párhuzamosan alapozható meg, s ezért nem mindegy, hogy milyen olvasástanítási módszert követünk. Olvasni mindegyikkel megtanul valahogy előbb-utóbb a gyerek, de írni nem.

Egy norvég főiskolai tanár, *Alfred Lie* részben megismételte, részben kiegészítette a fenti kísérletet, s hasonló eredményről számolt be "Elsőosztályos gyerekek szóanalizálási készségét fejlesztő program hatása" címen (1991). Meg erősítette Lundbergék azon megállapítását, hogy az eredmények később jobbak, az első és második év végén (a globális módszerekben ekkor szokott az összeroppanás bekövetkezni: a megnövekedett szóanyagot a gyerek memóriája már nem bírja, s a szükséges dekódolási mechanizmus még nincs kialakítva). Alfred Lie kiegészítése a következő: a szerző kétféle szóanalízist végeztetett: 1. a hangok egymás utáni felismertetését (ezt ő szekvenciás analízisnek nevezi), 2. a hangok felismerését a szó elején, közepén és végén (ezt pozíciós analízisnek nevezi). Az első, az egymás utáni analízis könnyebb volt a gyerekek számára. A szerző szerint különösen a gyengébb képességű gyerekek profitáltak sokat a tréningből. Természetesen, ezek a gyakorlatok is játékos módon voltak megoldva, sokat beszélgették a gyerekeket képekről. Lényegesnek tartották, hogy a gyerekek saját artikulációjukat figyeljék meg.

Egy dán pszichológus, *Mogens Hansen* is hasonló kísérleti programot szervezett, melynek neve *Project Cecilia* (1988), s melyről az IRA stockholmi 13. világkongresszusán tartott előadást 1990 júliusában. Egy koppenhágai külvárosi iskolában követtek mintegy 200 tanulót az óvodában és első osztályban. Főleg az óvodában foglalkoztak a nyelvi tudatosság fejlesztésével, első osztályban már inkább a strukturált szókincsfejlesztés és egyéb tevékenységek voltak a középpontban. Azt tételezték fel, hogy egy rendszeres tréning előkészíti a gyermekek olvasás-, írás- és matematikatanulását, vagyis pszichológiai terminussal élve – az olvasás jobb fejlődését és jobb feladattudatot vártak a kísérletben részt vevő gyerekektől (John Downing terminológiáját használja a szerző). A gyerekek lassabban, de gondosabban és kevesebb hibával olvastak, valóban jobban teljesítettek, mint a kontroll csoportban lévők.



Csak három kutatást emeltünk ki a számos közül. A nyelvi tudatosság kérdése valóban a figyelem középpontjában van jelenleg. Az alaposan megtervezett és gondosan lebonyolított skandináv kísérletek meggyőzően bizonyítják az előkészítés fontosságát. Meggyőzően bizonyítják azt is, hogy az olvasástanítás megalapozásának az időszaka a helyesírás-tanítás megalapozásának az időszaka is egyben.

3. *A tantárgytörténetből levonható tanulságok.* A metodika módszerei közé tartozik a kísérlet, a megfigyelés és a tapasztalatok átadása. Nemcsak a kortársak adhatják át tapasztalataikat, hanem a régiek is: ebben áll a tantárgytörténet ismeretének a jelentősége. A vitás kérdések eldöntéséhez, valamint a jövő tervezéséhez kötelességünk tudni azt, hogy régen milyen módszerek működtek, közülük melyek voltak sikeresek, és melyek voltak sikertelenek. Ezzel a céllal írtuk meg A magyar olvasástanítás története c. könyvet (Adamikné Jászó 1990).

Az, aki némiképp ismeri az olvasástanítás történetét, és ugyanakkor tanulmányozza a nyelvi tudatossággal kapcsolatos olvasáspszichológiai kutatásokat, felismeri, hogy a *fonológiai tréning* nem új találmány. *Előkészítés* néven működött, elsősorban a német oktatási hagyományokon építkező európai módszerekben, így nálunk is. Tudjuk például, hogy a Rein-féle herbartianus irányzat a múlt század hetvenes éveiben egy féléves (!) előkészítési szakaszt iktatott be a tulajdonképpeni olvasástanítás megkezdése elé (Rein 1878). A magyar olvasástanítás történetéből tudjuk – ezt a korszakot magam dolgoztam fel –, hogy az 1869-es tanterv a Gönczy Pál által kidolgozott írvaolvasztató módszert ajánlotta. Ez hangoztató, analitikus-szintetikus módszer volt, *hathetes előkészítési szakaszt iktatott be* az olvasástanítás elé, sőt a *beszéd- és értelemgyakorlatok* (később: beszélgetés) nevű tantárgy egyik feladata az olvasástanítás megalapozása volt.<sup>8</sup> Ez a gondos, előrelátó struktúra lényegében 1969-ig létezett. Ekkor, azaz az 1969-ben induló fővárosi kísérletben faragták le az előkészítő szakaszt, majd 1978-ban, az új olvasástanítási programok bevezetésekor eltörölték. Mindkét, 1978-ban bevezetett program, s később a Zsolnai-program is, szóképek olvasásával kezd. Teljesen megszüntették a részkészségek kimunkálását, s ezzel rengeteget ártottak nemcsak az olvasástanításnak, de a helyesírás tanításának is. A szótagolás kiiktatása (Ligeti 1974; Zsolnai – Zsolnai 1978) szintén felmérhetetlen károkat okozott, elsősorban a helyesírás tanításában.

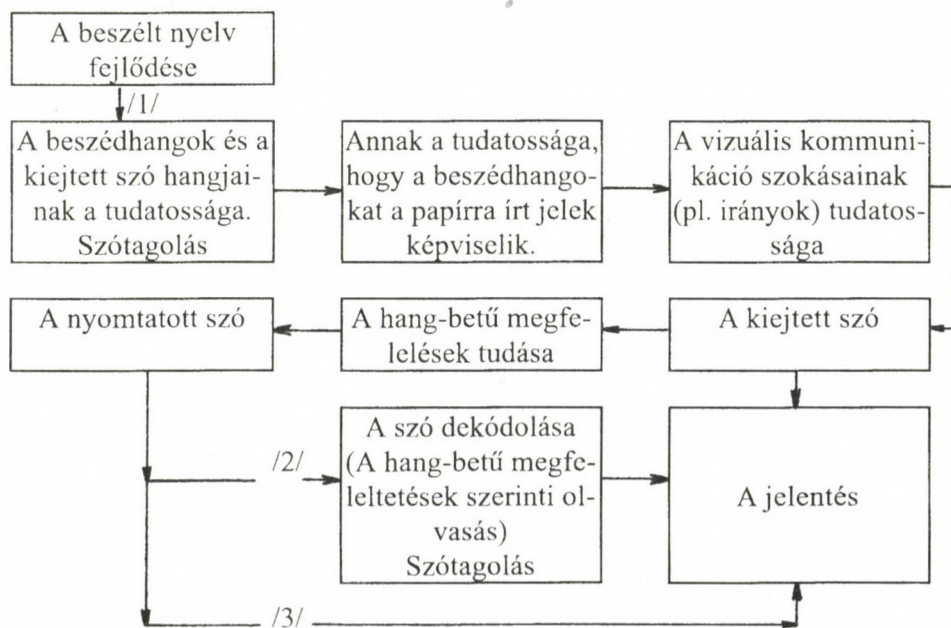
Az előkészítés kiiktatása, a részkészségek elhanyagolása az ún. diszlexiás és diszgráfias tanulók tömegeit produkálja, s ez is az oka annak, hogy az utóbbi évtizedben oly sokat romlott az írás és a helyesírás színvonala.

Érdekes, hogy Nyugat-Európában sem ismerik a tantárgytörténetet, maguk a németek is jelentettek meg olyan tanulmányt, melyben meg sem említik saját múlt századi módszereiket.<sup>10</sup> Az amerikai kutatók is lényegében saját jelenlegi gyakorlatukat ismerik, és nagy részük el sem tudja képzelni, hogy másképp is lehet tanítani, mint ahogyan azt mostanában teszik.



A gyakorlat gondjait és eredménytelenségét látva, a következőket fogalmazhatjuk meg: 1. semmiképpen sem szabad utánoznunk különféle külföldi módszereket; 2. viszont meg kell ismernünk őket és a külföldi elméleteket, kutatásokat; 3. ismernünk kell a tantárgy történetét; 4. mindezek segítségével egy olyan elméleti alapot kell teremteni, mely támogatja a gyakorlatot; 5. ily módon önállóvá kell tenni a módszertant, melynek segítője a pszichológia, és nem fordítva: nem szabad különféle pszichológiai kísérleteknek feláldozni a módszertant. Úgy látom, hogy az általam tanulmányozott olvasáspszichológiai kísérletek függenek attól, hogy milyen módon tanított gyerekanyagon végezték őket. Azok például akik azt állítják, hogy a nyelvi tudatosság megléte nem előfeltétele az olvasástanulásnak, s a nyelvi tudatosság az olvasástanulás során alakul ki, vajon nem csak globális módszerekkel tanított gyerekekkel találkoztak? Tudják-e vajon, hogy másképp is lehet tanítani? Tudják-e, hogy a 20. sz. kivételével mindig más volt a gyakorlat?

4. *Az olvasástanítás optimális modellje.* A Mattingly – Downing-féle elmélet alapján dolgozta ki az olvasástanítás modelljét egy angol szerző, *Douglas Pidgeon*. Tanulmányának címe: *Elmélet és gyakorlat az olvasástanulásban* (Pidgeon 1984). Ez modell a terminus eredeti értelmében, azt a folyamatot ábrázolja ugyanis, amely az olvasás tanulásakor az agyban végbemehet. Ugyanakkor megadja az olvasástanítás menetét is, a készségfejlesztés tervezését. Hazai körülményeinkhez alkalmaztam a modellt, bevettem ugyanis a számunkra oly fontos szótagolást, ami az eredetiből hiányzik.



A modellben a három szám a készség kialakulásának a három szakaszát jelenti. A 3.-kal jelölt útvonal az automatikus szakasz: ránézünk a nyomtatott szóra, és azonnal megértjük. A 2.-kal jelölt útvonal az elsajátítási szakasz, itt a nyomtatott szó egésze és a jelentés közé beiktatódik a dekódolás, vagyis a hangbetű megfelelések szerinti, valamint a szótagoló olvasás. Az 1.-vel jelölt útvonal a leghosszabb, itt történik az előkészítés. A tanítás kiindulása a beszélt nyelv, s mindvégig a kiejtett szó. Az előkészítés két lényeges pontja a szóalak felbontása szótagokra és hangokra, valamint az írás szokásainak a tudatosítása, pl. az irányok tanítása. Ennek a modellnek az "újdonsága" az előkészítés hangsúlyozása, valamint az auditív és a vizuális készségek egyensúlya, de mindig az auditív készségek (a beszélt nyelv, a kimondott szó) előtérbe helyezésével.

A globális módszer – valamint mindenféle globális előprogram azonnal a 3. útvonalat kívánja a gyerektől. Nincs az épület felépítve, s ezért omlik össze később. Mind a külföldi olvasáspszichológiai kísérletek (Downing, Lundberg stb.), mind a tantárgytörténeti tapasztalatok, mind a diszlexia problémája azt követeli, hogy az olvasási programba (ismét) beiktassuk az előkészítést, ügyeljünk a hangoztatás primátusára, s a gondos hang- és betűtanításra. Ebben a szellemben kívánok egy módszertant írni, mellyel remélhetőleg nemsokára elkészülök. Azt is nagyon remélem, hogy az ilyen szellemben tanított gyerekek írása és helyesírása is jó lesz.

### Jegyzetek

- <sup>1</sup> Az "A Dictionary of Reading and Related Terms" – IRA, 1981-c. szótárban pl. ez olvasható: deep structure meaning;...the meaning to which a spoken or written sentence refers which may not be apparent from its surface structure until verified by sentence transformation" (80).
- <sup>2</sup> Goodman, Yetta M.: (szerk.): How Children Construct Literacy. Piagetian Perspectives. IRA, 1990. Ebben a kötetben olasz, spanyol, brazil, izraeli és mexikói szerzők kísérleti beszámolóit olvashatók.
- <sup>3</sup> Például a Reading Teacher 1991. szeptemberi számában is van egy tanulmány, mely azzal foglalkozik, hogyan kell a tanulókat értékelni az egész nyelvi módszert alkalmazó osztályokban.
- <sup>4</sup> A nyelvi tudatosság fogalmára többféle terminus van használatban: metalinguistics, meta-linguistic skills, linguistic awareness, language awareness, szűkebb értelemben: lexical awareness, phonological awareness stb.
- <sup>5</sup> A helyesírás morfofonetikus jellege az angolra is érvényes, vö.: Venetzky, R.: The Structure of English Orthography. The Hague, 1970.
- <sup>6</sup> A Downing – Valtin által szerkesztett 1984-es könyvben rövidebben, Olvasáspszichológiájukban egy hosszú fejezetben foglalkoznak az olvasással mint készséggel.
- <sup>7</sup> Downing, I., Ayers, D. és Schaefer, B., 1983. Ez az ún. LARR-teszt.
- <sup>8</sup> A beszélgetést 1963-ban szüntették meg, a környezetismeretet hozták létre helyette.
- <sup>9</sup> A helyesírás romlásáról a következő megfigyeléseket tettem. Immár 17. éve annak, hogy rendszeresen felmérjük a tanítóképzőbe jövő hallgatók helyesírásának a színvonalát. Sok olyan hibát követnek el, melyek mögött grammatikai tudatlanság húzódik meg (pl. a játssza, hagyja, készség stb. helyesírása). Újabban jelentkeznek az elválasztási hibák, a zöngésedés elrontása,



tehát olyan problémák, melyek régebben még gimnáziumban sem jelentkeztek. Feltűnő az ly-j és a t végű igék felszólító módjának a begyakorlatlansága. Az esetek többségét szótagolással be lehetne gyakorolni.

- <sup>10</sup> A Reading Teacher 1991. augusztusi számában olvashattuk Franz Biglmaier beszámolóját a német olvasástanítás történetéről. Olyan dolgokról sem ír, melyeket a múlt században nálunk is tudott minden tanító.

### Irodalom

- Adamikné Jászó Anna (szerk.): A magyar olvasástanítás története. Budapest 1990.
- Bolla Kálmán: A leíró hangtan vázlata. In: Fejezetek a magyar leíró hangtanból. Budapest 1982.
- Chomsky, C.: Stages in language development and reading exposure. Harvard Educational Review. 42, 1972, 1–33.
- Downing, J. (szerk.): Comparative Reading. New York 1973; Downing, J. – Leong, C. K.: Psychology of Reading. New York 1982.
- Downing, J. – Valtin, R.: Language Awareness and Learning to Read. Language and Communication 17. 1984.
- Downing, J. – Ayers, D. – Schaefer, B.: Linguistic Awareness in Reading Readiness. Windsor, England 1983.
- Ehri, L. C.: How orthography alters spoken language competencies in children learning to read and spell. In: Downing, J. – Valtin, R. (szerk.): Language Awareness and Learning to Read. New York 1984, 119–147. A továbbiakban Downing – Valtin, 1984.
- Fekete László: Magyar kiejtési szótár. Budapest 1992.
- Goodman, K. S.: Reading: A Psycholinguistic Guessing Game. In: Singer, H. – Ruddell, R. B. (szerk.): Theoretical Models and Processes of reading. IRA, 1970.
- Goodman, Y. M.: (szerk.): How Children Construct Literacy. Piagetian Perspectives. IRA, 1990.
- Gósy Mária: Beszédszlelés. Budapest 1989a.
- Gósy Mária: GMP Beszédszlelési és beszédmegértési teljesítmény. Tesztcsomag. Budapest 1989b.
- Gósy Mária: Óvónői beszédviselkedés. Budapest 1990.
- Hangay Zoltán: Nyelvészeti munka a Budapesti Tanítóképző Főiskolán. Magyar Nyelvőr 112, 1988, 67–70.
- Hansen, M.: Project Cecilia: A Language Enrichment Programme for 5–7 Year-old Children. International Research in Reading. Volume 5. No.1, 1988, 44–66.
- Liberman, A. M. – Cooper, F. S. – Shankweiler, D. P. – Studdert-Kennedy, M.: Perception of the Speech Code. Psychological Review, 74, 1967, 431–461.
- Lie, A.: Effect of a Training Program for Stimulating Skills in Word Analysis in First-Grade Children. Reading Research Quarterly XXV, 3. 1991, 234–250.
- Ligeti Róbert: Az olvasás tanítása szintetikus és globális programokkal. Tanító 2. 1974, 6–8.
- Lundberg, I. – Frost, L. – Petersen, O.: Effects of an Extensive Program for Stimulating Phonological Awareness in Preschool Children. Reading Research Quarterly XXIII, 3. 1988, 263–284.
- Mattingly, I. G.: Reading, linguistic awareness, and language acquisition. In: Downing – Valtin, 1984, 9–25.
- Mattingly, I. G.: Reading, the Linguistic Process, and Linguistic Awareness. In: Kavanagh, J. F. – Mattingly, I. G. (szerk.): Language by Ear and by Eye. Cambridge, Mass. 1972.

- Melyiket válasszam? OPI-MM. Budapest 1990.
- Nagy J. József (szerk.): Anyanyelvi tantárgypedagógia. Budapest 1983.
- Pidgeon, D.: Theory and Practice in Learning to Read. In: Downing – Valtin, 1984, 173–191.
- Rein, W. – Pickel, A. – Scheller, E.: Das erste Schuljahr. Eisenach 1878.
- Szikszai Lajosné: A transzformációs elemzés felhasználása a tanítóképző főiskolákon folyó nyelvtanításban. Magyar Nyelvőr 112. 1988, 207–218.
- Szikszai Lajosné: Strukturális eljárások a tanítóképző főiskola nyelvtanításában. Magyar Nyelvőr 113. 1989, 340–347.
- Vernon, N. D.: Backwardness in Reading. Cambridge, England 1957.
- Zsolnai József – Zsolnai Józsefné: Az olvasástanítási módszerek összehasonlító értékelése. Tanító 6–7. 1978, 46–53.











297---